Differential Equations

James S. Cook
Liberty University
Department of Mathematics

2025 and beyond

preface

format of my notes

These notes were prepared with LaTeX. You'll notice a number of standard conventions in my notes:

1. definitions are in green.
2. remarks are in red.
3. theorems, propositions, lemmas and corollaries are in blue.
4. proofs start with a **Proof:** and are concluded with a □. However, we also use the discuss... theorem format where a calculation/discussion leads to a theorem and the formal proof is left to the reader.

The purpose of these notes is to organize a large part of the theory for this course. Your text has much additional discussion about how to calculate given problems and, more importantly, the detailed analysis of many applied problems. I include some applications in these notes, but my focus is much more on the topic of computation and, where possible, the extent of our knowledge. There are some comments in the text and in my notes which are not central to the course. To learn what is most important you should come to class every time we meet.

sources and philosophy of my notes

I draw from a number of excellent sources to create these notes. Naturally, having taught from the text for about a dozen courses, Nagle Saff and Snider has had great influence in my thinking on DEqns, however, more recently I have been reading: the classic *Introduction to Differential Equations* by Albert Rabenstein. I reccomend that text for further reading. In particular, Rabenstein's treatment of existence and convergence is far deeper than I attempt in these notes. In addition, the text *Differential Equations with Applications* by Ritger and Rose is a classic text with many details that are lost in the more recent generation of texts. I am also planning to consult the texts by Rice & Strange, Zill, Edwards & Penny , Finney and Ostberg, Coddington, Arnold, Zachmanoglou and Thoe, Hille, Ince, Martin, Campbell as well as others I'll add to this list once these notes are more complete (some year).

Additional examples are also posted. My website has several hundred pages of solutions from problems in Nagle Saff and Snider. I hope you will read these notes and the required text as you study differential equations this semester. My old lecture notes are sometimes useful, but I hope the theory in these notes is superior in clarity and extent. My primary goal is the **algebraic** justification of the computational essentials for differential equations. The organization of the required text is doubtless not quite the same as these notes. I have recommended certain homeworks to help you grow past the central theme of my teaching.

I should mention, topics which are very incomplete in this current version of my notes are:

1. theory of ODEs, Picard iteration (here is the most offensive offense as it regards math majors taking this course)

2. the special case of Frobenius method (Nagle Saff and Snider pick up the slack in what is left out here)
3. the magic formulas of Ritger and Rose on constant coefficient case (these appear in the older texts like Ince, but have been lost in many of the newer more expensive but less useful texts)
4. theory of orthogonal functions, Fourier techniques
5. separation of variables to solve PDEs (this could be another half course if expanded properly)
6. linear system analysis via Greens functions and the transfer function (this is another course if expanded properly)

I have several hundred pages of handwritten notes and solutions posted on the course webpage. There are discussions and calculations in some of those notes which extend past what I provide here. Naturally, the best way to know what is important in your course is to attend class, attempt all the assignments and do your best on the tests. The only way to really learn DEqns is to calculate on your own. Remember, watching me solve DEqns is not the same is doing them yourself. Practice is key.

Last semester I spent many hours collecting all the homework I have assigned in Differential Equations over the years. I called these *Practice Problems*. I've place these at the end of chapters in this version of the notes. If you look over the course website http://www.supermath.info/DEqns.html then you'll find many such problems solved.

James Cook, July 13, 2025.

version 4.01

In the Spring 2026 Semester I am working on reformulating the notes. I'll try to announce my progress as the semester unfolds.
James Cook, January 19, 2026

version 5.01

The Spring 2026 Semester is complete. The notes now have a more or less complete treatment of PDEs at the level of this course. It could afford some physically motivating additions in a future edit. Also, a Chapter on Variational Calculus is now added. It can be lectured in a week as a gentle introduction to the calculus of variations with an emphasis on its application to classical mechanics. I should credit Annabeth Hughes for pointing out over 100 errors in the previous version of these notes.

James Cook, May 16, 2026

version 6.7.

Contents

Chapter 1

terminology and goals

If you are not interested in terminology, scope and general philosophy then you may skip ahead to Chapter 2 where we begin our study of how to solve differential equations. The current chapter is primarily concerned with explaining the big picture of differential equations. In particular, we briefly introduce terminology here and we attempt to give a good set of examples which show what it means to solve a differential eqaution. However, we do not explain here (for the most part) how to find such solutions. Examples of techinques on how to find solutions are given in later chapters.

1.1 terms and conditions

A **differential equation** (or DEqn) is simply an equation which involves derivatives. The **order** of a differential equation is the highest derivative which appears nontrivially in the DEqn. The **domain of definition** is the set of points for which the expression defining the DEqn exists. We consider real independent variables and for the most part real dependent variables, however we will have occasion to consider complex-valued objects. The complexity will occur in the range but not in the domain. We continue to use the usual notations for derivatives and integrals in this course. I will not define these here, but we should all understand the meaning of the symbols below: the following are examples of **ordinary dervatives**

$$\frac{dy}{dx} = y' \qquad \frac{d^2y}{dt^2} = y'' \qquad \frac{d^ny}{dt^n} = y^{(n)} \qquad \frac{d\vec{r}}{dt} = \left\langle \frac{dx}{dt}, \frac{dy}{dt} \right\rangle$$

because the **dependent variables** depend on only one **independent variable**. Notation is not reserved globally in this course. Sometimes x is an independent variable whereas other times it is used as a dependent variable, context is key; $\frac{dy}{dx}$ suggests x is independent and y is dependent whereas $\frac{dx}{dt}$ has independent variable t and dependent variable x. A DEqn which involves only ordinary derivatives is called an **Ordinary Differential Equation** or as is often customary an "ODE". The majority of this course we focus our efforts on solving and analyzing ODEs. However, even in the most basic first order differential equations the concept of partial differentiation and functions of several variables play a key and notable role. For example, an **n-th order ODE** is an equation of the form $F(y^{(n)}, y^{(n-1)}, \dots, y'', y', y, x) = 0$. For example,

$$y'' + 3y' + 4y^2 = 0 \;\; (n = 2) \qquad y^{(k)}(x) - y^2 - xy = 0 \;\; (n = k)$$

When $n = 1$ we say we have a **first-order ODE**, often it is convenient to write such an ODE in the form $\frac{dy}{dx} = f(x, y)$. For example,

$$\frac{dy}{dx} = x^2 + y^2 \text{ has } f(x,y) = x^2 + y^2 \qquad \frac{dr}{d\theta} = r\theta + 7 \text{ has } f(r,\theta) = r\theta + 7$$

A **system of ODEs** is a set of ODEs which share a common independent variable and a set of several dependent variables. For example, the following system has dependent variables x, y, z and independent variable t:

$$\frac{dx}{dt} = x^2 + y + \sin(t)z, \qquad \frac{d^2y}{dt^2} = xyz + e^t, \qquad \frac{dz}{dt} = \sqrt{x^2 + y^2 + z^2}.$$

The examples given up to this point were all **nonlinear** ODEs because the dependent variable or it's derivatives appeared in a nonlinear manner. Such equations are actually quite challenging to solve and the general theory is not found in introductory textbooks. It turns out that we can solve many nonlinear first order ODEs. In contrast, solvable higher-order nonlinear problems are for the most part beyond the reach of this course.

A **n-th order linear ODE in standard form** is a DEqn of the form:

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y + a_o y = g$$

where $a_n, a_{n-1}, \dots, a_1, a_o$ are the **coefficients** which are generally functions and g is the **forcing function** or **inhomogenous term**. Continuing, if an n-th order ODE has $g = 0$ then we say it is a **homogenous** DEqn. When the coefficients are simply constants then the DEqn is said to be a **constant coefficient** DEqn. It turns out that we can solve **any** constant coefficient n-th order ODE. A system of ODEs for which each DEqn is linear is called a **system of linear DEqns**. For example:

$$x'' = x + y + z + t \qquad y'' = x - y + 2z, \qquad z'' = z + t^3.$$

If each linear DEqn in the system has constant coefficients then the system is also said to be a **constant coefficient** system of linear ODEs. We will see how to solve any **constant coefficient** linear system. Linear differential equations with nonconstant coefficients are not as simple to solve, however, we will solve a number of interesting problems via the series technique.

Partial derivatives are defined for functions or variables which depend on multiple independent variables. For example,

$$u_x = \frac{\partial u}{\partial x} \qquad T_{xy} = \frac{\partial^2 T}{\partial y \partial x} \qquad \nabla^2 \Phi = \partial_x^2 \Phi + \partial_y^2 \Phi + \partial_z^2 \Phi \qquad \nabla \bullet \vec{E} \qquad \nabla \times \vec{B}.$$

You should have studied the divergence $\nabla \bullet \vec{E}$ and curl $\nabla \times \vec{B}$ in multivariable calculus. The expression $\nabla^2 \Phi$ is called the **Laplacian** of Φ. A DEqn which involves partial derivatives is called a **Partial Differential Equation** or as is often customary a "PDE". We study PDEs towards the conclusion of this course. It turns out that solving PDEs is naturally accomplished by a mixture of ODE and general series techniques.

1.2 philosophy and goals

What is our primary goal in this course? In a nutshell; to find the solution of a differential equation. Obviously this begs a question: "what is the solution to a DEqn?" I would answer that as follows:

> **a solution to a DEqn is a function or level set for which the given differential equation is a differential consequence of the solution.**[1]

In other words, a solution to a given DEqn is some object that satifies the DEqn when you "plug it in". I mean *differential consequence* to mean equation(s) produced through differentiation of the given function(s) or level set(s). I've added (s) to reflect the fact that some differential equations require multiple functions or level sets and some also require more than one differentiation.

Example 1.2.1. *$y = \cos(x)$ is a solution of $y''+y = 0$ since $y = \cos(x)$ implies $y'' = -\cos(x) = -y$ thus $y''+y = 0$. By the same token, if $x' = \frac{dx}{dt}$ then $x''+x = 0$ has solution $x = \cos t$. This solution models the motion of a mass moved by a spring without friction.*

In the example above we saw the solution is a function. This is not always possible in practice.

Example 1.2.2. *If you implicitly differentiate $xy^3 + y^2 = \sin(x) + 3$ then it is easy to see the equation $xy^3+y^2 = \sin(x)+3$ defines a solution of $y^3+3xy^2\frac{dy}{dx}+2y\frac{dy}{dx}-\cos(x) = 0$. I would rather not find the solution as a function of x in this example.*

Example 1.2.3. *If $F(x, y) = x^2+y^3 = 10$ then $dF = 2xdx+3y^2dy = 0$ is a differential consequence of the equation $F(x, y) = 10$.*

Example 1.2.4. *Let $\vec{F}(x, y, z) = \langle x, 2y, 3z\rangle$ and consider $\vec{r}(t) = \langle x_oe^t, y_oe^{2t}, z_oe^{3t}\rangle$ which is to say parametrically, $x = x_oe^t$, $y = y_oe^{2t}$ and $z = z_oe^{3t}$. Calculate $\frac{d\vec{r}}{dt} = \langle x_oe^t, 2y_oe^{2t}, 3z_oe^{3t}\rangle = \vec{F}(\vec{r}(t))$. Thus $\vec{r}(t) = \langle x_oe^t, y_oe^{2t}, z_oe^{3t}\rangle$ solves $\frac{d\vec{r}}{dt} = \vec{F}(\vec{r}(t))$ with initial condition $\vec{r}(0) = (x_o, y_o, z_o)$. Geometrically, $\vec{r}$ gives the* **integral curve** *or* **flowline** *of the vector field $\vec{F}$.*

In every example I gave thus far I found a differential equation by differentiating a given object. I have yet to show any method for how we can **find** the solution when given a differential equation. These examples merely intend to illustrate what is meant by the term **solution**. Let's give a precise definition for a particular type of differential equation:

Definition 1.2.5. *explicit solution for an n-th order ODE:*

> It is convenient to define an **explicit solution** on $I \subseteq \mathbb{R}$ for an n-th order ODE $F(y^{(n)}, y^{(n-1)}, \dots, y'', y', y, x) = 0$ is a function ϕ such that $F(\phi^{(n)}(x), \phi^{(n-1)}(x), \dots, \phi''(x), \phi'(x), \phi(x), x) = 0$ for all $x \in I$.

In many problems we do not discuss I since $I = \mathbb{R}$ and it is obvious (for instance, Example 1.2.1), however, when we discuss singular points in the later portion of the course the domain of definition plays an interesting and non-trivial role.

Very well, the concept of a solution is not too difficult. Let's ask a harder question: how do we find solutions? Begin with a simple problem:

$$\frac{dy}{dx} = 0 \quad \Rightarrow \quad \int \frac{dy}{dx}dx = \int 0\, dx \quad \Rightarrow \quad \boxed{y = c_o.}$$

[1]in some semesters of Math 332 we study differential equations in differential forms, you can picture differential equations as exterior differential systems where the solutions are submanifolds which correspond naturally to the vanishing of a differential form. It gives you a coordinate free formulation of differential equations.

Integrating revealed that solutions of $y' = 0$ are simply constant functions. Notice each distinct value for c_o yields a distinct solution. What about $y'' = 0$?

$$\frac{d^2y}{dx^2} = 0 \quad \Rightarrow \quad \int \frac{d}{dx}\left[\frac{dy}{dx}\right] dx = \int 0\, dx \quad \Rightarrow \quad \frac{dy}{dx} = c_1.$$

Integrate indefinitely once more,

$$\int \frac{dy}{dx}\, dx = \int c_1\, dx \quad \Rightarrow \quad \boxed{y = c_1 x + c_o.}$$

Therefore, we find a whole family of solutions for $y'' = 0$; the solution set of $y'' = 0$ is

$$\{f \mid f(x) = c_1 x + c_o \text{ for } c_o, c_1 \in \mathbb{R}\} = span\{1, x\}$$

where I use the notation **span** to indicate the set of all linear combinations. Students of linear algebra will recognize the solution set is a two-dimensional subspace of function space. Continuing in this pattern, to solve $y^{(n)}(x) = 0$ we can integrate n-times to derive

$$\boxed{y = \frac{1}{(n-1)!} c_{n-1} x^{n-1} + \cdots + \frac{1}{2} c_2 x^2 + c_1 x + c_o.}$$

The solution set here can be written as $span\{1, x, \dots, x^{n-1}\}$ and we note it forms an n-dimensional subspace of function space[2] We should know from Taylor's Theorem in second semester calculus the constants are given by $y^{(n)}(0) = c_n$ since the solution is a Taylor polynomial centered at $x = 0$. Hence we can write the solution in terms of the value of y, y', y'' etc... at $x = 0$: suppose $y^{(n)}(0) = y_n$ are given **initial conditions** then

$$\boxed{y(x) = \frac{1}{(n-1)!} y_n x^{n-1} + \cdots + \frac{1}{2} y_2 x^2 + y_1 x + y_o.}$$

We see that the arbitary constants we derived allow for different initial conditions which are possible. In calculus we add C to the indefinite integral to allow for all possible antiderivatives. In truth, $\int f(x)\, dx = \{F \mid F'(x) = f(x)\}$, it is a set of antiderivatives of the integrand f. However, almost nobody writes the set-notation because it is quite cumbersome. Likewise, in our current context we will be looking for the solution set of a DEqn, but we will call it the **general solution**. The **general solution** is usually many solutions which are indexed by a few arbitrary constants.

Example 1.2.6. *For example, the general solution to* $y''' - 4y'' + 3y' = 0$ *is* $y = c_1 + c_2 e^t + c_3 e^{3t}$.

Example 1.2.7. *Or the general solution to* $x' = -y$ *and* $y' = x$ *is given by* $x = c_1 \cos(t) + c_2 \sin(t)$ *and* $y = c_1 \sin(t) - c_2 \cos(t)$.

To be careful, it is not always the case there is a whole family of solutions (as in the two previous examples), there are curious DEqns which have just one solution or even none. Consider the following example showcasing two **nonlinear** ODEs:

[2]This sentence is here for those students who have taken Linear Algebra and know what I mean by basis and dimension of a vector space. Linear differential ODEs are greatly aided by linear algebra since the homogeneous solution set is a finite dimensional subspace of function space. In short, this means the algebra of linear ODEs is actually pretty easy.

Example 1.2.8. *Notice that $y = 0$ is the only solution of*

$$(y')^2 + y^2 = 0$$

whereas

$$(y')^2 + y^2 = -1$$

has no solutions.

There are other nonlinear examples where the constants index over most of the solution set, but miss a few special solutions.

We just saw that integration can sometimes solve a problem. However, can we always integrate? I mentioned that $y'' + y = 0$ has solution $y = \cos(x)$. In fact, you can show that $y = c_1\cos(x) + c_2\sin(x)$ is the general solution. Does integration reveal this directly?

$$y'' = -y \quad \Rightarrow \quad \int y''\,dx = \int -y\,dx \quad \Rightarrow \quad y' = C + \int -y\,dx$$

at this point we're stuck. In order to integrate we need to know the formula for y. But, that is what we are trying to find! DEqns that allow for solution by direct integration are somewhat rare.

Example 1.2.9. *Suppose $y''(x) = g(x)$ for some continuous function g has a solution which is obtained from twice integrating the DEQn: I'll find a solution in terms of the initial condtions at $x = 0$:*

$$y'' = g \quad \Rightarrow \quad \int_0^x y''(t)\,dt = \int_0^x g(t)\,dt \quad \Rightarrow \quad y'(x) = y'(0) + \int_0^x g(t)\,dt$$

integrate once more, this time use s as the dummy variable of integration,
note $y'(s) = y'(0) + \int_0^s g(t)\,dt$ hence

$$\int_0^x y'(s)\,ds = \int_0^x \left[y'(0) + \int_0^s g(t)\,dt\right] ds \quad \Rightarrow \quad \boxed{y(x) = y(0) + y'(0)x + \int_0^x \int_0^s g(t)\,dt\,ds.}$$

Note that the integral above does not involve y itself, if we were give a nice enough function g then we might be able to find an simple form of the solution in terms of elementary functions.

If direct integration is not how to solve *all* DEqns then what should we do? Well, that's what we intend to learn this semester. Overall it is very similar to second semester calculus and integration. We make educated guesses then we differentiate to check if it worked. Once we find something that works then we look for ways to reformulate a broader class of problems back into those basic templates. But, the key here is guessing. Not blind guessing though. Often we make a general guess that has flexibility built-in via a few *parameters*. If the guess or *ansatz* is wise then the parameters are naturally chosen through some condition derived from the given DEqn.

If we make a guess then how do we know we didn't miss some possibility? I suppose we don't know. Unless we discuss some of the theory of differential equations. Fortunately there are deep and broad existence theorems which not only say the problems we are trying to solve are solvable, even more, the theory tells us how many *linearly independent* solutions we must find. The theory has the most to say about the linear case. However, as you can see from the nonlinear examples $(y')^2 + y^2 = 0$ and $(y')^2 + y^2 = -1$, there is not much we can easily say in general about the structure of solutions for nonlinear ODEs.

We say a set of conditions are **initial conditions (IC)** if they are all given at the same value of an independent variable. In contrast, **boundary conditions** or **BCs** are given at two or more values of the independent variables. If pair a DEqn with a set of initial conditions then the problem of solving the DEqn subject to the intial conditions is called an **initial value problem** or **IVP**. If pair a DEqn with a set of boundary conditions then the problem of solving the DEqn subject to the boundary conditions is called a **boundary value problem** or **BVP**. For example,

$$y'' + y = 0 \qquad \text{with } y(0) = 0 \text{ and } y'(0) = 1$$

is an IVP. The **unique** solution is simply $y(x) = \sin(x)$. On the other hand,

$$y'' + y = 0 \qquad \text{with } y(0) = 0 \text{ and } y(\pi) = 0$$

is a BVP which has a family of solutions $y(x) = \sin(nx)$ indexed by $n \in \mathbb{N}$. Other BVPs may have no solutions at all. We study BVPs in our analysis of PDEs towards the end of this course. Given a linear ODE with continuous coefficient and forcing functions on $I \subseteq \mathbb{R}$ the IVP has a unique solution which extends to all of I. In particular, the constant coefficient linear ODE has solutions on $\mathbb{R}$. This is a very nice result which is physically natural; given a DEqn which models some phenomenon we find that the same thing happens every time we start the system with a particular initial condition. In contrast, nonlinear DEqns sometime allow for the same initial condition to yield infinitely many possible solutions.

The majority of our efforts are placed on finding functions or equations which give solutions to DEqns. These are **quantitative** results. There is also much that can be said **qualitatively** or even **graphically**. In particular, we can study **autonomous** systems of the form $dx/dt = f(x, y)$ and $dy/dt = g(x, y)$ by plotting the **direction field** of the system. The solutions can be seen by tracing out curves which line-up with the arrows in the direction field. Software[3] will plot direction fields for autonomous systems and you can easily see what types of behaviour are possible. All of this is possible when explicit quantitative solutions are intractable.

Numerical solutions to DEqns is one topic these notes neglect. If you are interested in numerical methods then you should try to take the numerical methods course. That course very useful to those who go on to business and industry, probably linear algebra is the only other course we offer which has as wide an applicability. The other topic which is neglected in these notes is rigor in the sense of mathematical analysis. I have no intention of belaboring issues of existence and convergence for most discussions and I tend to be very terse about domains except where they matter.

I should mention, we use the concept of a differential in this course. Recall that if F is a function of $x_1, x_2, \dots, x_n$ then we defined

$$dF = \frac{\partial F}{\partial x_1}dx_1 + \frac{\partial F}{\partial x_2}dx_2 + \cdots + \frac{\partial F}{\partial x_n}dx_n = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}dx_i.$$

When the symbol d acts on an equation it is understood we are taking the total differential. I assume that is reasonable to either write

$$\frac{dy}{dx} = f(x, y) \qquad \text{or} \qquad dy = f(x, y)dx \qquad \text{or} \qquad f(x, y)dx - dy = 0.$$

[3]such as pplane of **Matlab** which is built-in to an applet linked on the course page

These expressions are meaningful and I explain in great detail in the advanced calculus course how the method of differentials enjoys its sucess on the back of the implicit and inverse function theorems. However, this is not advanced calculus so I will not prove or deeply discuss those things here. I am just going to use them, *formally* if you wish. More serious is our lack of focus on existence and convergence, those analytical discussions tend to beg questions from real analysis and are a bit beyond the level of these notes and this course. That said let us say a little bit about a nuanced issue of the domain of definition for a differential equation.

Example 1.2.10. *Consider the differential equation* $\frac{dy}{dx} = \frac{x}{y}$*. Clearly* $y = 0$ *is not a solution since we cannot divide by zero. In contrast,* $ydy = xdx$ *has solution* $y = 0$*. This should not be surprising, we know from algebra that changing the form of an expression may change its domain. In some sense, the differential equation* $ydy - xdx = 0$ *is more general than* $\frac{dy}{dx} = \frac{x}{y}$ *(which misses the solution* $y = 0$*) or* $\frac{dx}{dy} = \frac{y}{x}$ *(which misses the solution* $x = 0$*).*

1.3 a short overview of differential equations in basic physics

I'll speak to what I know a little about. These comments are for the reductionists in the audience.

- **Newtonian Mechanics** is based on Newton's Second Law which is stated in terms of a time derivative of three functions. We use vector notation to say it succinctly as

$$\boxed{\frac{d\vec{P}}{dt} = \vec{F}_{net}}$$

  where $\vec{P}$ is the momentum and $\vec{F}_{net}$ is the force applied.

- **Lagrangian Mechanics** is the proper way of stating Newtonian mechanics. It centers its focus on energy and conserved quantities. It is mathematically equivalent to Newtonian Mechanics for some systems. The fundamental equations are called the Euler Lagrange equations they follow from Hamilton's principle of least action $\delta S = \delta \int L \; dt = 0$,

$$\boxed{\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{y}_j}\right] = \frac{\partial L}{\partial y_j}}$$

  where y_j is a generalized coordinate. We can write down Euler Lagrange equation for one or many generalized coordinates are appropriate to the problem. Lagrangian mechanics allows you to derive equations of physics in all sorts of curvy geometries. Geometric constraints are easily implemented by Lagrange multipliers. In any event, the mathematics here is integration, differentiation and to see the big picture variational calculus[4]

- **Electricity and Magnetism** boils down to solving Maxwell's equations subject to various boundary conditions:

$$\boxed{\nabla \cdot \vec{B} = 0, \qquad \nabla \cdot \vec{E} = \frac{\rho}{\epsilon_o}, \qquad \nabla \times \vec{B} = \mu_o \vec{J} + \mu_o \epsilon_o \frac{\partial \vec{E}}{\partial t} \qquad \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}.}$$

  Again, the mathematics here is calculus of several variables and vector notations. In other words, the mathematics of electromagnetism is vector calculus.

[4] I sometimes cover variational calculus in the Advanced Calculus course Math 332.

- **Special Relativity** also uses vector calculus. However, linear algebra is really needed to properly understand the general structure of Lorentz transformations. Mathematically this is actually not so far removed from electromagnetism. In fact, electromagnetism as discovered by Maxwell[5] around 1860 naturally included Einstein's special relativity. In relativitic coordinate free differential form language Maxwell's equations are simply stated as

$$\boxed{dF = 0, \qquad d * F = *J.}$$

  Newtonian mechanics is inconsistent with these equations thus Einstein's theory was inevitable. I should mention, we can also derive Maxwell's Equations from a variational calculus approach where the Euler Lagrange equations are given for the 4-potential on spacetime.

- **General Relativity** uses calculus on manifolds. A manifold is a curved surface which allows for calculus in local coordinates. The geometry of the manifold encodes the influence of gravity and conversely the presence of mass curves space and time. Einstein's field equations are nonlinear partial differential equations for which only a few special cases have well-known solutions. For example, the blackhole solution stemming from Schwarzchild's treatment of the spherically symmetric case is actually not too difficult to derive. Many more realistic scenarios require sophisticated computer simulation. Indeed, much of current astrophysics is more or less a game of minecraft which is taken far more seriously than perhaps we ought[6].

- **Quantum Mechanics** based on Schrodinger's equation which is a partial differential equation (much like Maxwell's equations) governing a complex wave function. Alternatively, quantum mechanics can be formulated through the path integral formalism as championed by Richard Feynman.

- **Quantum Field Theory** is used to frame modern physics. The mathematics is not entirely understood. However, Lie groups, Lie algebras, supermanifolds, jet-bundles, algebraic geometry are likely to be part of the correct mathematical context. Physicists will say this is done, but mathematicians do not in general agree. To understand quantum field theory one needs to master calculus, differential equations and more generally develop an ability to conquer very long calculations.

In fact, all modern technical fields in one way or another have differential equations at their core. This is why you are expected to take this course.

> If a system has variables where the change in one variable or a parameter causes a change in one or more of the variables then it is likely that a differential equation can be used to express the rule(s) governing the change.

Differential equations are also used to model phenomenon which are not basic; population models, radioactive decay, chemical reactions, mixing tank problems, heating and cooling, financial markets, fluid flow, a snowball which gathers snow as it falls, a bus stopping as it rolls through a giant vat of peanut butter, a rope falling off a table etc... the list is endless. If you think about the course you took in physics you'll realize that you were asked about specific times and events, but there is

[5]It should be mentioned, that Maxwell did not use vector calculus notation, instead his equations were initially expressed in quaternionic terms. Oliver Heaviside reformulated Maxwell's Equations into a notation more like the usual vector calculus if I understand the history correctly, although, I have not checked too deeply on this claim

[6]although, my children do take minecraft very seriously at times

also the question of how the objects move once the forces start to act. The step-by-step continuous picture of the motion is going to be the solution to the differential equation called Newton's Second Law. Beyond the specific examples we look at in this course, it is my hope you gain a more general appreciation of the method. In a nutshell, the leap in concept is to use derivatives to model things.

1.4 course overview

- **Chapter 1:** you've almost read the whole thing. By now you should realize it is to be read once now and once again at the end of the semester.
- **Chapter 2:** we study first order ODEs. One way or another it usually comes back to some sort of integration.
- **Chapter 3:** we study n-th order linear ODEs. I'll lecture on a method presented in Ritger and Rose which appears magical, however, we don't just want answers. We want understanding and this is brought to us from a powerful new way of thinking called the *operator method*. We'll see how many nontrivial problems are reduced to algebra. Variation of parameters takes care of the rest.
- **Chapter 4:** we study systems of linear ODEs, we'll need matrices and vectors to properly treat this topic. The concept of eigenvectors and eigenvalues plays an important role, however the operator method shines bright once more here.
- **Chapter 5:** energy analysis and the phase plane approach. In other chapters our goal has almost always to find a solution , but here we study properties of the solution without actually finding it. This qualitative approach can reveal much without too much effort. When paired with the convenient pplane software we can ascertain many things with a minimum of effort.
- **Chapter 6:** the method of Laplace transforms is shown to solve problems with discontinuous, even infinite, forcing functions with ease.
- **Chapter 7:** some problems are too tricky for the method of Chapter 3. We are forced to resort to power series techniques. Moreover, some problems escape power series as well. The method of Frobenius helps us capture behaviour near regular singular points. The functions discovered here have tremendous application across the sciences.
- **Chapter 8:** homogeneous boundary value problems are studied. We study heat, wave and Laplace's equations via a mixture of our the BVP solutions and the Fourier technique.

Chapter 2

ordinary first order problem

We wish to solve problems of the form $\frac{dy}{dx} = f(x,y)$. An **explicit solution** is a function $\phi : I \subset \mathbb{R} \to \mathbb{R}$ such that $\frac{d\phi}{dx} = f(x, \phi(x))$ for all $x \in I$. There are several common techniques to solve such problems, although, in general the solution may be impossible to find in terms of elementary functions. You should already anticipate this fact from second semester calculus. Performing an indefinite integration is equivalent to solving a differential equation; observe that

$$\int e^{x^2} dx = y \qquad \Leftrightarrow \qquad \frac{dy}{dx} = e^{x^2}.$$

you may recall that the integration above is not amenable to elementary techniques[1]. However, it is simple enough to solve the problem with series techniques. Using term-by-term integration,

$$\int e^{x^2}\, dx = \int \sum_{n=0}^{\infty} \frac{(x^2)^n}{n!}\, dx = \sum_{n=0}^{\infty} \frac{1}{n!} \int x^{2n}\, dx = \sum_{n=0}^{\infty} \frac{1}{n!(2n+1)} x^{2n+1} + c.$$

This simple calculation shows that $y = x + \frac{1}{6}x^3 + \frac{1}{10}x^5 + \cdots$ will solve $\frac{dy}{dx} = e^{x^2}$. We will return to the application of series techniques to find analytic solutions later in this course. For this chapter, we wish to discuss those techniques which allow us to solve first order problems via algebra and integrals of elementary functions. There are really three[2] main techniques:

1. separation of variables
2. integrating factor method
3. identification of problem as an exact equation

Beyond that we study substitutions which bring the problem back to one of the three problems above in a new set of variables. The methods of this chapter are by no means complete or algorithmic. Solving arbitrary first order problems is an art, not unlike the problem of parametrizing a level curve. That said, it is not a hidden art, it is one we all must master.

[1] the proof of that is not elementary!

[2] you could divide these differently, it is true that the integrating factor technique is just a special substitution

2.1 separation of variables

Suppose you are faced with the problem $\frac{dy}{dx} = f(x,y)$. If it happens that f can be factored into a product of functions $f(x,y) = g(x)h(y)$ then the problem is said to be **separable**. Proceed formally for now, suppose $h(y) \neq 0$,

$$\frac{dy}{dx} = g(x)h(y) \quad \Rightarrow \quad \frac{dy}{h(y)} = g(x)\,dx \quad \Rightarrow \quad \int \frac{dy}{h(y)} = \int g(x)\,dx$$

Ideally, we can perform the integrations above and solve for y to find an explicit solution. However, it may even be preferrable to not solve for y and capture the solution in an **implicit** form. Let me provide a couple examples before I prove the method at the end of this section.

Example 2.1.1. Problem: *Solve* $\frac{dy}{dx} = 2xy$.

Solution: *Separate variables to find* $\int \frac{dy}{y} = \int 2x\,dx$ *hence* $\ln|y| = x^2 + c$. *Exponentiate to obtain* $|y| = e^{x^2+c} = e^c e^{x^2}$. *The constant* $e^c \neq 0$ *however, the absolute value allows for either* $\pm$. *Moreover, we can also observe directly that* $y = 0$ *solves the problem. We find* $\boxed{y = ke^{x^2}}$ *is the general solution to the problem.*

An explicit solution of the differential equation is like an antiderivative of a given integrand. The general solution is like the indefinite integral of a given integrand. The general solution and the indefinite integral are not functions, instead, they are a family of functions of which each is an explicit solution or an antiderivative. Notice that for the problem of indefinite integration the constant can always just be thoughtlessly tacked on at the end and that will nicely index over all the possible antiderivatives. On the other hand, for a differential equation the constant could appear in many other ways.

Example 2.1.2. Problem: *Solve* $\frac{dy}{dx} = \frac{-2x}{2y}$.

Solution: *separate variables and find* $\int 2y\,dy = -\int 2x\,dx$ *hence* $y^2 = -x^2 + c$. *We find* $x^2 + y^2 = c$. *It is clear that* $c < 0$ *give no interesting solutions. Therefore, without loss of generality, we assume* $c \geq 0$ *and denote* $c = R^2$ *where* $R \geq 0$. *Altogether we find* $\boxed{x^2 + y^2 = R^2}$ *is the general* ***implicit*** *solution to the problem. To find an explicit solution we need to focus our efforts, there are two cases:*

1. *if* (a,b) *is a point on the solution and* $b > 0$ *then* $y = \sqrt{a^2 + b^2 - x^2}$.
2. *if* (a,b) *is a point on the solution and* $b < 0$ *then* $y = -\sqrt{a^2 + b^2 - x^2}$.

Notice here the constant appeared inside the square-root. I find the implicit formulation of the solution the most natural for the example above, it is obvious we have circles of radius R. To capture a single circle we need two function graphs. Generally, given an implicit solution we can solve for an explicit solution locally. The implicit function theorems of advanced calculus give explicit conditions on when this is possible.

Example 2.1.3. Problem: *Solve* $\frac{dy}{dx} = e^{x + 2\ln|y|}$.

Solution: *recall* $e^{x+\ln|y|^2} = e^x e^{\ln|y|^2} = e^x |y|^2 = e^x y^2$. *Separate variables in view of this algebra:*

$$\frac{dy}{y^2} = e^x\,dx \quad \Rightarrow \quad \frac{-1}{y} = e^x + C \quad \Rightarrow \quad \boxed{y = \frac{-1}{e^x + C}}.$$

When I began this section I mentioned the justification was *formal*. I meant that to indicate the calculation seems plausible, but it is not justified. We now show that the method is in fact justified. In short, I show that the notation works.

Proposition 2.1.4. *separation of variables:*

The differential equation $\frac{dy}{dx} = g(x)h(y)$ has an implicit solution given by

$$\int \frac{dy}{h(y)} = \int g(x)\,dx$$

for (x, y) such that $h(y) \neq 0$.

Proof: to say the integrals above are an implicit solution to $\frac{dy}{dx} = g(x)h(y)$ means that the differential equation is a differential consequence of the integral equation. In other words, if we differentiate the integral equation we should hope to recover the given DEqn. Let's see how this happens, differentiate implicitly,

$$\frac{d}{dx}\int \frac{dy}{h(y)} = \frac{d}{dx}\int g(x)\,dx \quad \Rightarrow \quad \frac{1}{h(y)}\frac{dy}{dx} = g(x) \quad \Rightarrow \quad \frac{dy}{dx} = h(y)g(x). \qquad \square$$

Remark 2.1.5.

Technically, there is a gap in the proof above. How did I know implicit differentiation was possible? Is it clear that the integral equation defines y as a function of x at least locally? We could use the implicit function theorem on the level curve $F(x, y) = \int \frac{dy}{h(y)} - \int g(x)\,dx = 0$. Observe that $\frac{\partial F}{\partial y} = \frac{1}{h(y)} \neq 0$ hence the implicit function theorem provides the existence of a function ϕ which has $F(x, \phi(x)) = 0$ at points near the given point with $h(y) \neq 0$. This comment comes to you from the advanced calculus course.

2.1.1 geometric applications

Let us take a break from introducing new methods for a moment and study a pair of applications. To be clear, in general the problems I introduce here require other techniques to solve. I merely focus on examples here where the math works out nice and easy.

Differential equations make quick work of certain geometry problems. We introduce two such problems in this section:

- **Integral Curves of $\langle P, Q\rangle$:** If $\vec{r} = \langle x, y\rangle$ is an integral curve of $\vec{F}$ then we have that $\vec{F}(\vec{r}(t)) = \frac{d\vec{r}}{dt}$. This is what we mean by an integral curve; it is a curve whose tangent field aligns with the given vector field. In other words, the integral curve is a **streamline** of the vector field. Notice $\frac{d\vec{r}}{dt} = \left\langle \frac{dx}{dt}, \frac{dy}{dt} \right\rangle$ thus we must solve

$$\frac{dx}{dt} = P \qquad \& \qquad \frac{dy}{dt} = Q$$

  This is a **system of first order ODEs** and depending on the details of P and Q we may or may not be able to explicitly solve such a problem. Later in the course we learn how to solve linear systems of ODEs. That said, we can use a little calculus trick to eliminate time t:

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{Q}{P}.$$

Therefore, to find the Cartesian equation of the integral curves to $\vec{F} = \langle P, Q\rangle$ we simply must solve $\frac{dy}{dx} = \frac{Q}{P}$.

- **Orthogonal Trajectories:** To find the orthogonal trajectory of $F(x,y) = c$ we use calculus to determine $\frac{dy}{dx} = S(x,y)$ for the given curve(s) then solve $\frac{dy}{dx} = -1/S(x,y)$. The resulting family of solutions will intersect $F(x,y) = c$ orthogonally.

Let's expand these bullet points with explicit examples.

Example 2.1.6. Problem: *Find the Cartesian equation of the integral curves of* $\vec{F} = \langle x, y\rangle$.

Solution: *Identify* $Q = y$ *and* $P = x$ *thus integral curves are solutions of* $\frac{dy}{dx} = \frac{y}{x}$. *We can solve via separation of variables:*

$$\frac{dy}{dx} = \frac{y}{x} \;\Rightarrow\; \int \frac{dy}{y} = \int \frac{dx}{x} \;\Rightarrow\; \ln|y| = \ln|x| + c \;\Rightarrow\; |y| = e^{\ln|x|+c} = e^c|x| \;\Rightarrow\; \boxed{y = mx.}$$

This makes sense, the given vector field points radially out from the origin.

Example 2.1.7. Problem: *Find the Cartesian equation of the integral curves of* $\vec{F} = \langle -y, x\rangle$.

Solution: *Identify* $Q = x$ *and* $P = -y$ *thus integral curves are solutions of* $\frac{dy}{dx} = \frac{x}{-y}$. *We can solve via separation of variables:*

$$\frac{dy}{dx} = \frac{x}{-y} \;\Rightarrow\; \int y\,dy = -\int x\,dx \;\Rightarrow\; y^2/2 = -x^2/2 + c \;\Rightarrow\; x^2 + y^2 = -2c \;\Rightarrow\; \boxed{x^2 + y^2 = R^2}.$$

I set $-2c = R^2$ *since it is clear we need* $-2c \geq 0$ *for the equation to have a non-empty solution. Of course, the integral curves here are concentric circles about the origin. If you sketch the vector field it becomes clear this solution is to be expected.*

Example 2.1.8. Problem: *Find the orthogonal trajectories to the curve* $x^2 + y^3 = 8$.

Solution: *if* $x^2 + y^3 = 8$ *then* $2x + 3y^2\frac{dy}{dx} = 0$ *hence* $\frac{dy}{dx} = \frac{-2x}{3y^2}$. *Identify* $S(x,y) = \frac{-2x}{3y^2}$ *hence* $\frac{-1}{S} = \frac{3y^2}{2x}$. *To find orthogonal trajectories we must solve* $\frac{dy}{dx} = \frac{3y^2}{2x}$. *We use separation of variables once more:*

$$\frac{dy}{dx} = \frac{3y^2}{2x} \;\Rightarrow\; \int \frac{dy}{y^2} = \frac{3}{2}\int \frac{dx}{x} \;\Rightarrow\; \frac{-1}{y} = \frac{3}{2}\ln|x| + c \;\Rightarrow\; \boxed{y = \frac{-1}{\ln|x|^{3/2} + c}.}$$

Example 2.1.9. Problem: *Find the orthogonal trajectories to the curve* $y = k/x$ *where* k *is constant.*

Solution: *if* $y = k/x$ *then* $xy = k$ *thus* $y + x\frac{dy}{dx} = 0$ *hence* $\frac{dy}{dx} = \frac{-y}{x}$. *Identify* $S(x,y) = \frac{-y}{x}$ *hence* $\frac{-1}{S} = \frac{-1}{\frac{-y}{x}} = \frac{x}{y}$. *To find orthogonal trajectories we must solve* $\frac{dy}{dx} = \frac{x}{y}$. *We use separation of variables:*

$$\frac{dy}{dx} = \frac{x}{y} \;\Rightarrow\; \int y\,dy = \int x\,dx \;\Rightarrow\; y^2/2 = x^2/2 + c \;\Rightarrow\; \boxed{y^2 - x^2 = 2c.}$$

The orthogonal trajectories to $y = k/x$ *are the family of hyperbolas given by* $y^2 - x^2 = 2c$.

2.2 integrating factor method

Let p and q be continuous functions. The following differential equation is called a **linear differential equation** in standard form:

$$\boxed{\frac{dy}{dx} + py = q \qquad (\star)}$$

Our goal in this section is to solve equations of this type. Fortunately, linear differential equations are very nice and the solution exists and is not too hard to find in general, well, at least up-to a few integrations.

Notice, we cannot directly separate variables because of the py term. A natural thing to notice is that it sort of looks like a product, maybe if we multiplied by some new function I then we could separate and integrate: multiply $\star$ by I,

$$I\frac{dy}{dx} + pIy = qI$$

Now, if we choose I such that $\frac{dI}{dx} = pI$ then the equation above separates by the product rule:

$$\frac{dI}{dx} = pI \quad \Rightarrow \quad I\frac{dy}{dx} + \frac{dI}{dx}y = qI \quad \Rightarrow \quad \frac{d}{dx}\big[Iy\big] = qI \quad \Rightarrow \quad Iy = \int qI\,dx \quad \Rightarrow \quad \boxed{y = \frac{1}{I}\int qI\,dx.}$$

Very well, but, is it possible to find such a function I? Can we solve $\frac{dI}{dx} = pI$? Yes. Separate variables,

$$\frac{dI}{dx} = pI \quad \Rightarrow \quad \frac{dI}{I} = p\,dx \quad \Rightarrow \quad \ln(I) = \int p\,dx \quad \Rightarrow \quad \boxed{I = e^{\int p\,dx}.}$$

Proposition 2.2.1. *integrating factor method:*

Suppose p, q are continuous functions which define the linear differential equation $\frac{dy}{dx}+py = q$ (label this $\star$). We can solve $\star$ by the following algorithm:

(1.) define $I = exp(\int p\,dx)$,

(2.) multiply $\star$ by I,

(3.) apply the product rule to write $I\star$ as $\frac{d}{dx}\big[Iy\big] = Iq$.

(4.) integrate both sides,

(5.) find general solution $y = \frac{1}{I}\int Iq\,dx$.

Proof: Define $I = e^{\int p\,dx}$, note that p is continuous thus the antiderivative of p exists by the FTC. Calculate,

$$\frac{dI}{dx} = \frac{d}{dx}e^{\int p\,dx} = e^{\int p\,dx}\frac{d}{dx}\int p\,dx = pe^{\int p\,dx} = pI.$$

Multiply $\star$ by I, use calculation above, and apply the product rule:

$$I\frac{dy}{dx} + Ipy = Iq \quad \Rightarrow \quad I\frac{dy}{dx} + \frac{dI}{dx}y = Iq \quad \Rightarrow \quad \frac{d}{dx}\big[Iy\big] = Iq.$$

Integrate both sides,

$$\int \frac{d}{dx}\big[Iy\big]\,dx = \int Iq\,dx \quad \Rightarrow \quad Iy = \int Iq\,dx \quad \Rightarrow \quad y = \frac{1}{I}\int Iq\,dx. \qquad \square$$

The integration in $y = \frac{1}{I}\int Iq\,dx$ is indefinite. It follows that we could write $y = \frac{C}{I} + \frac{1}{I}\int Iq\,dx$. Note once more that the constant is not simply added to the solution.

Example 2.2.2. Problem: *find the general solution of* $\frac{dy}{dx} + \frac{2}{x}y = 3$

Solution: *Identify that* $p = 2/x$ *for this linear DE. Calculate, for* $x \neq 0$,

$$I = exp\Big(\int \frac{2dx}{x}\Big) = exp(2\ln|x|) = exp(\ln|x|^2) = |x|^2 = x^2$$

Multiply the DEqn by $I = x^2$ *and then apply the reverse product rule;*

$$x^2\frac{dy}{dx} + 2xy = 3x^2 \quad \Rightarrow \quad \frac{d}{dx}\big[\,x^2y\,\big] = 3x^2$$

Integrate both sides to obtain $x^2y = x^3 + c$ *therefore* $\boxed{y = x + c/x^2}$.

We could also write $y(x) = x + c/x^2$ to emphasize that we have determined y as a function of x.

Example 2.2.3. Problem: *let* r *be a real constant and suppose* g *is a continuous function, find the general solution of* $\frac{dy}{dt} - ry = g$

Solution: *Identify that* $p = -r$ *for this linear DE with independent variable* t. *Calculate,*

$$I = exp\Big(\int -r\,dt\Big) = e^{-rt}$$

Multiply the DEqn by $I = e^{-rt}$ *and then apply the reverse product rule;*

$$e^{-rt}\frac{dy}{dt} - re^{-rt}y = ge^{-rt} \quad \Rightarrow \quad \frac{d}{dt}\big[\,e^{-rt}y\,\big] = ge^{-rt}$$

Integrate both sides to obtain $e^{-rt}y = \int g(t)e^{-rt}\,dt + c$ *therefore* $\boxed{y(t) = ce^{rt} + e^{rt}\int g(t)e^{-rt}\,dt}$.

Now that we worked this in general it's fun to look at a few special cases:

(1.) if $g = 0$ *then* $\boxed{y(t) = ce^{-rt}}$.

(2.) if $g(t) = e^{rt}$ *then* $y(t) = ce^{rt} + e^{rt}\int e^{-rt}e^{rt}\,dt$ *hence* $\boxed{y(t) = ce^{rt} + te^{rt}}$.

(3.) if $r \neq s$ *and* $g(t) = e^{st}$ *then* $y(t) = ce^{rt} + e^{rt}\int e^{st}e^{-rt}\,dt = ce^{rt} + e^{rt}\int e^{(s-r)t}\,dt$ *consqeuently we find that* $y(t) = ce^{rt} + \frac{1}{s-r}e^{rt}e^{(s-r)t}$ *and thus* $\boxed{y(t) = ce^{rt} + \dfrac{1}{s-r}e^{st}}$.

It might be interesting to examine $s \to r$ *in (3.) above.*

2.3 exact equations

Before we discuss the theory I need to introduce some new notation:

Definition 2.3.1. *Pffafian form of a differential equation*

Let M, N be functions of x, y then $Mdx + Ndy = 0$ is a differential equation in **Pfaffian form**.

For example, if $\frac{dy}{dx} = f(x,y)$ then $dy - f(x,y)dx = 0$ is the differential equation in its Pfaffian form. One advantage of the Pfaffian form is that it puts x, y on an equal footing. There is no artificial requirement that y be a function of x implicit within the set-up, instead x and y appear in the same way. The natural solution to a differential equation in Pfaffian form is a level curve.

Example 2.3.2. *Consider the circle* $x^2 + y^2 = R^2$ *note that* $2xdx + 2ydy = 0$ *hence the circle is a solution curve of* $2xdx + 2ydy = 0$

Recall the total differential[3] of a function $F : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ was defined by:

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy.$$

Let k be a constant and observe that $F(x,y) = k$ has $dF = 0dx + 0dy = 0$. Conversely, if we are given $\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0$ then we find natural solutions of the form $F(x,y) = k$ for appropriate constants k. Let us summarize the technique:

Proposition 2.3.3. *exact equations:*

If differential equation $Mdx + Ndy = 0$ has $M = \frac{\partial F}{\partial x}$ and $N = \frac{\partial F}{\partial y}$ for some differentiable function F then the solutions to the differential equation are given by the level-curves of F.

A **level-curve** of F is simply the collection of points (x, y) which solve $F(x,y) = k$ for a constant k. You could also call the solution set of $F(x,y) = k$ the k-level curve of F or the fiber $F^{-1}\{k\}$.

Remark 2.3.4.

Consider vector field $\langle M, N\rangle$ and consider the solution of $Mdx+Ndy = 0$ given by $F(x,y) = c$. Suppose $t \mapsto \vec{r}(t)$ parametrizes the solution $F(x,y) = c$ then $F(\vec{r}(t)) = c$. Observe $\nabla F = \langle \partial_x F, \partial_y F\rangle = \langle M, N\rangle$. Notice the chain-rule gives:

$$\frac{d}{dt}(c) = \frac{d}{dt}\left[F(\vec{r}(t))\right] = \nabla F \bullet \frac{d\vec{r}}{dt} = \langle M, N\rangle \bullet \left\langle \frac{dx}{dt}, \frac{dy}{dt}\right\rangle$$

Of course, $\frac{d}{dt}(c) = 0$ hence we find the tangent vector $\left\langle \frac{dx}{dt}, \frac{dy}{dt}\right\rangle$ to the level curve $F(x,y) = c$ is perpendicular to the vector $\langle M, N\rangle$ if we study a particular point on the level curve $F(x,y) = c$. Geometrically, we should expect the solutions of $Mdx + Ndy = 0$ gives us a family of level curves which perpendicularly intercept the streamlines of $\langle M, N\rangle$.

[3] You might wonder what precisely dx and dy mean in such a context. If you want to really know then take advanced calculus. For our purposes here is suffices to inform you that you can multiply and divide by differentials, these formal algebraic operations are in fact a short-hand for deeper arguments justified by the implicit and/or inverse function theorems. But, again, that's advanced calculus.

Example 2.3.5. **Problem:** *find the solutions of* $y^2dx + 2xydy = 0$.

Solution: *we wish to find* F *such that*

$$\frac{\partial F}{\partial x} = y^2 \qquad \& \qquad \frac{\partial F}{\partial y} = 2xy$$

You can integrate these equations holding the non-integrated variable fixed,

$$\frac{\partial F}{\partial x} = y^2 \quad \Rightarrow \quad F(x,y) = \int y^2\, dx = xy^2 + C_1(y)$$

$$\frac{\partial F}{\partial y} = 2xy \quad \Rightarrow \quad F(x,y) = \int 2xy\, dy = xy^2 + C_2(x)$$

It follows that $F(x,y) = xy^2$ *should suffice. Indeed a short calculation shows that the given differential equation in just* $dF = 0$ *hence the solutions have the form* $xy^2 = k$. *One special solution is* $x = 0$ *and* y *free, this is allowed by the given differential equation, but sometimes you might not count this a solution. You can also find the explicit solutions here without too much trouble:* $y^2 = k/x$ *hence* $y = \pm\sqrt{k/x}$. *These solutions foliate the plane into disjoint families in the four quadrants:*

$$k > 0 \text{ and } + \text{ in I}, \qquad k < 0 \text{ and } + \text{ in II}, \qquad k < 0 \text{ and } - \text{ in III}, \qquad k > 0 \text{ and } - \text{ in IV}$$

The coordinate axes separate these cases and are themselves rather special solutions for the given DEqn.

The explicit integration to find F is not really necessary if you can make an educated guess. That is the approach I adopt for most problems.

Example 2.3.6. **Problem:** *find the solutions of* $2xy^2dx + (2x^2y - \sin(y))dy = 0$

Solution: *observe that the function* $F(x,y) = x^2y^2 + \cos(y)$ *has*

$$\frac{\partial F}{\partial x} = 2xy^2 \qquad \& \qquad \frac{\partial F}{\partial y} = 2x^2y - \sin(y)$$

Consequently, the given differential equation is nothing more than $dF = 0$ *which has obvious solutions of the form* $x^2y^2 + \cos(y) = k$.

I invite the reader to find explicit local solutions for this problem. I think I'll stick with the level curve view-point for examples like this one.

Example 2.3.7. **Problem:** *find the solutions of* $\frac{xdx+ydy}{x^2+y^2} = 0$

Solution: *observe that the function* $F(x,y) = \frac{1}{2}\ln(x^2+y^2)$ *has*

$$\frac{\partial F}{\partial x} = \frac{x}{x^2+y^2} \qquad \& \qquad \frac{\partial F}{\partial y} = \frac{y}{x^2+y^2}$$

Consequently, the given differential equation is nothing more than $dF = 0$ *which has curious solutions of the form* $\frac{1}{2}\ln(x^2+y^2) = k$. *If you expoentiate this equation it yields* $\sqrt{x^2+y^2} = e^k$. *We can see that the unit-circle corresponds to* $k = 0$ *whereas generally the* k*-level curve has radius* e^k.

Notice that $2x + 2y = 0$ and $\frac{xdx+ydy}{x^2+y^2} = 0$ share nearly the same set of solutions. The origin is the only thing which distinguishes these examples. This raises a question we should think about. When are two differential equations equivalent? I would offer this definition: two differential equations are equivalent if they share the same solution set. This is the natural extension of the concept we already know from algebra. Naturally the next question to ask is: how can we modify a given differential equation to obtain an equivalent differential equation? This is something we have to think about as the course progresses. Whenever we perform some operation to a differential equation we ought to ask, *did I just change the solution set?*. For example, multiplying $2x+2y = 0$ by $\frac{1}{2(x^2+y^2)}$ removed the origin from the solution set of $\frac{xdx+ydy}{x^2+y^2} = 0$.

2.3.1 conservative vector fields and exact equations

You should recognize the search for F in the examples above from an analgous problem in multivariable calculus[4] Suppose $\vec{G} = \langle M, N\rangle$ is conservative on U with potential function F such that $\vec{G} = \nabla F$. Pick a point (x_o, y_o) and let C be the level curve of F which starts at (x_o, y_o)[6]. Recall that the tangent vector field of the level curve $F(x, y) = k$ is perpendicular to the gradient vector field ∇F along C. It follows that $\int_C \nabla F \bullet d\vec{r} = 0$. Or, in the differential notation for line-integrals, $\int_C M dx + N dy = 0$.

Continuing our discussion, suppose (x_1, y_1) is the endpoint of C. Let us define the line-segment L_1 from (x_o, y_o) to (x_o, y_1) and the line-segment L_2 from (x_o, y_1) to (x_1, y_1). The curve $L_1 \cup L_2$ connects (x_o, y_o) to (x_1, y_1). By path-independence of conservative vector fields we know that $\int_{L_1\cup L_2} \vec{G}\bullet d\vec{r} = \int_C \vec{G}\bullet d\vec{r}$. It follows that[7]:

$$\begin{aligned} 0 = \int_{L_1\cup L_2} \vec{G}\bullet d\vec{r} &= \int_{L_1} N\,dy + \int_{L_2} M\,dx \\ &= \int_{y_o}^{y_1} N(x_o, t)\,dt + \int_{x_o}^{x_1} M(t, y_1)\,dt \end{aligned}$$

Let $x_1 = x$ and $y_1 = y$ and observe that the equation

$$0 = \int_{y_o}^{y} N(x_o, t)\,dt + \int_{x_o}^{x} M(t, y)\,dt$$

ought to provide the level-curve solution of the exact equation $Mdx + Ndy = 0$ which passes through the point (x_o, y_o). For future reference let me summarize our discussion here:

---

[4] Let us briefly review the results we derived for conservative vector fields in multivariable calculus. Recall that $\vec{G} = \langle M, N\rangle$ is conservative iff there exists a potential function[5] F such that $\vec{G} = \nabla F = \langle \partial_x F, \partial_y F\rangle$ on $dom(\vec{G})$. Furthermore, it is known that $\vec{G} = \langle M, N\rangle$ is conservative on a simply connected domain iff $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ for all points in the domain. The Fundamental Theorem of Calculus for line-integrals states if C is a curve from P to Q then $\int_C \nabla F \bullet d\vec{r} = F(Q) - F(P)$. It follows that conservative vector fields have the property of path-independence. In particular, if $\vec{G}$ is conservative on U and C_1, C_2 are paths beginning and ending at the same points then $\int_{C_1} \vec{G}\bullet d\vec{r} = \int_{C_2} \vec{G}\bullet d\vec{r}$.

[6] the level curve extends past C, we just want to make (x_o, y_o) the starting point

[7] Recall that if C is parametrized by $\vec{r}(t) = \langle x(t), y(t)\rangle$ for $t_1 \leq t \leq t_2$ then the line-integral of $\vec{G} = \langle M, N\rangle$ is by definition:

$$\int_C M dx + N dy = \int_{t_1}^{t_2} \left[M(x(t), y(t))\frac{dx}{dt} + N(x(t), y(t))\frac{dy}{dt}\right] dt$$

I implicitly make use of this definition in the derivation that follows.

Proposition 2.3.8. *solution by line-integral for exact equations:*

Suppose the differential equation $Mdx + Ndy = 0$ is exact on a simply connected region U then the solution through $(x_o, y_o) \in U$ is given implicitly by

$$\int_{x_o}^{x} M(t, y)\, dt + \int_{y_o}^{y} N(x_o, t)\, dt = 0.$$

Perhaps you doubt this result. We can check it by taking the total differential of the proposed solution:

$$\begin{aligned} d\left[\int_{x_o}^{x} M(t,y)\,dt\right] &= \frac{\partial}{\partial x}\left[\int_{x_o}^{x} M(t,y)\,dt\right]dx + \frac{\partial}{\partial y}\left[\int_{x_o}^{x} M(t,y)\,dt\right]dy \\ &= M(x,y)dx + \left[\int_{x_o}^{x} \frac{\partial M}{\partial y}(t,y)\,dt\right]dy \\ &= M(x,y)dx + \left[\int_{x_o}^{x} \frac{\partial N}{\partial x}(t,y)\,dt\right]dy \qquad \text{since } \partial_x N = \partial_y M \\ &= M(x,y)dx + \big[N(x,y) - N(x_o,y)\big]dy \end{aligned}$$

On the other hand,

$$d\left[\int_{y_o}^{y} N(x_o,t)\,dt\right] = \frac{\partial}{\partial x}\left[\int_{y_o}^{y} N(x_o,t)\,dt\right]dx + \frac{\partial}{\partial y}\left[\int_{y_o}^{y} N(x_o,t)\,dt\right]dy = N(x_o,y)dy$$

Add the above results together to see that $M(x,y)dx + N(x,y)dy = 0$ is a differential consequence of the proposed solution. In other words, it works.

Example 2.3.9. Problem: *find the solutions of* $(2xy + e^y)dx + (2y + x^2 + e^y)dy = 0$ *through* $(0,0)$.

Solution: *note* $M(x,y) = 2xy + e^y$ *and* $N(x,y) = 2y + x^2 + e^y$ *has* $\partial_y M = \partial_x N$. *Apply Proposition 2.3.8*

$$\begin{aligned} \int_0^x M(t,y)\,dt + \int_0^y N(0,t)\,dt = 0 \quad &\Rightarrow \quad \int_0^x (2ty + e^y)\,dt + \int_0^y (2t + e^t)\,dt = 0 \\ &\Rightarrow \quad \Big(t^2 y + te^y\Big)\Big|_0^x + \Big(t^2 + e^t\Big)\Big|_0^y = 0 \\ &\Rightarrow \quad \boxed{x^2 y + xe^y + y^2 + e^y - 1 = 0.} \end{aligned}$$

You can easily verify that $(0,0)$ *is a point on the curve boxed above.*

The technique illustrated in the example above is missing from many differential equations texts, I happened to discover it in the excellent text by Ritger and Rose *Differential Equations with Applications*. I suppose the real power of Proposition 2.3.8 is to capture formulas for an arbitrary point with a minimum of calculation:

Example 2.3.10. Problem: *find the solutions of $2x\,dx + 2y\,dy = 0$ through (x_o, y_o).*

Solution: *note $M(x,y) = 2x$ and $N(x,y) = 2y$ has $\partial_y M = \partial_x N$. Apply Proposition 2.3.8*

$$\begin{aligned}
\int_{x_o}^{x} M(t,y)\,dt + \int_{y_o}^{y} N(x_o,t)\,dt = 0 \quad &\Rightarrow \quad \int_{x_o}^{x} 2t\,dt + \int_{y_o}^{y} 2t\,dt = 0 \\
&\Rightarrow \quad t^2\Big|_{x_o}^{x} + t^2\Big|_{y_o}^{y} = 0 \\
&\Rightarrow \quad x^2 - x_o^2 + y^2 - y_o^2 = 0. \\
&\Rightarrow \quad \boxed{x^2 + y^2 = x_o^2 + y_o^2.}
\end{aligned}$$

The solutions are circles with radius $\sqrt{x_o^2 + y_o^2}$.

You can solve exact equations without Proposition 2.3.8, but I like how this result ties the math back to multivariable calculus.

Example 2.3.11. Problem: *find the solutions of $E_1\,dx + E_2\,dy = 0$ through (x_o, y_o). Assume $\partial_x E_2 = \partial_y E_1$.*

Solution: *the derivatiion of Proposition 2.3.8 showed that the solution of $E_1\,dx + E_2\,dy = 0$ is given by level curves of the potential function for $\vec{E} = \langle E_1, E_2 \rangle$. In particular, if $\vec{E} = -\nabla V$, where the minus is customary in physics, then the solution is simply given by the equipotential curve $V(x,y) = V(x_o, y_o)$. In other words, we could interpret the examples in terms of voltage and electric fields. That is an important, real-world, application of this mathematics.*

Remark 2.3.12.

In physics we sometimes define the electric potential by $V(\vec{r}) = -\int_{\mathcal{O}}^{\vec{r}} \vec{E} \bullet d\vec{r}$ this formulation of the potential implies $-\nabla V = \vec{E}$. On the other hand, if we think of this two-dimensionally then $\vec{E} \bullet d\vec{r} = E_1 dx + E_2 dy$ and so $E_1 dx + E_2 dy = -dV$ hence $V(\vec{r}) = c$ defines a solution of $E_1 dx + E_2 dy = 0$. We call such a level curve an **equipotential**.

Likewise, a conservative vector field $\vec{F} = \langle F_1, F_2 \rangle = -\nabla U$ where U is the potential energy. We could define $U(\vec{r}) = -\int_{\mathcal{O}}^{\vec{r}} \vec{F} \bullet d\vec{r}$ (here $\mathcal{O}$ defines the zero for the potential energy). If U is defined by this integral then $-\nabla U = \vec{F}$. In terms of differential equations, $F_1 dx + F_2 dy = 0$ has solution $U(x,y) = c$, a curve is a curve of constant potential energy.

2.3.2 Green's Theorem and the closed condition

Given a differential equation in Pfaffian form $Pdx + Qdy = 0$ if we are given that $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$ over a simply connected subset U of the plane then Green's Theorem applies to $D \subseteq U$ and we find:

$$\int_{\partial D} Pdx + Qdy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = 0$$

Thus $\langle P, Q \rangle$ is a conservative vector field on U since ∂D constitutes an arbitrary loop within U. You might thing this means ∂D is a solution of $Pdx + Qdy = 0$, but notice we only know the integral of

$Pdx + Qdy$ is zero around all of ∂D. In particular, $Pdx + Qdy$ can be both positive and negative on ∂D in such a way that it cancels out in the continuous sum around ∂D. That said, since $\langle P, Q\rangle$ is conservative on U we know there exists $F : U \to \mathbb{R}$ for which $\nabla F = \langle \partial_x F, \partial_y F\rangle = \langle P, Q\rangle$ hence

$$dF = Pdx + Qdy$$

and it becomes clear that $F(x, y) = c$ serves to define the solution since $dF = dc = 0$ for the points on the level curve defined by $F(x, y) = c$.

Suppose $Mdx + Ndy = 0$ is a differential equation which may or may not be exact. Suppose there exists F for which $dF = Mdx + Ndy$ then

$$\frac{\partial F}{\partial x} = M \qquad \& \qquad \frac{\partial F}{\partial y} = N \;\;\Rightarrow\;\; \frac{\partial M}{\partial y} = \frac{\partial}{\partial y}\left[\frac{\partial F}{\partial x}\right] = \frac{\partial}{\partial x}\left[\frac{\partial F}{\partial y}\right] = \frac{\partial N}{\partial x}$$

We say $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ is the **closed condition** for the differential equation $Mdx + Ndy = 0$. The calculation here shows that **if** the differential equation $Mdx+Ndy = 0$ is exact **then** the coefficient functions M, N must solve the closed condition $M_y = N_x$. Conversely, if the closed condition $M_y = N_x$ holds for some subset $U \subseteq \mathbb{R}^2$ then it is not generally true that $Mdx + Ndy = 0$ is exact. However, if we add the stipulation that the closed condition holds on a simply connected subset of $\mathbb{R}^2$ then it is true that the closed condition implies the differential equation is exact. The proof of this assertion is given at the start of this section. Let me collect the claims of this section for future reference:

Proposition 2.3.13. *closed condition for $Mdx + Ndy = 0$:*

> If $\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$ for some point $p \in U \subseteq \mathbb{R}^2$ then $Mdx + Ndy = 0$ is not an exact equation on U. However, if U is a simply connected subset of $\mathbb{R}^2$ for which $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ for each point in U then $Mdx + Ndy = 0$ is an exact equation.

Remember, to say $Mdx + Ndy = 0$ is an exact equations means that we can solve $M = \frac{\partial F}{\partial x}$ and $N = \frac{\partial F}{\partial y}$ simultaneously. Forgive me if I repeat myself a bit in the next section.

2.3.3 inexact equations and integrating factors

Consider once more the Pfaffian form $Mdx+Ndy = 0$. If $\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$ at some point P then we cannot find a potential function for a set which contains P. It follows that we can state the following no-go proposition for the problem of exact equations.

Proposition 2.3.14. *inexact equations:*

> If differential equation $Mdx + Ndy = 0$ has $\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$ then $Mdx + Ndy = 0$ is **inexact**. In other words, if $\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$ then there does not exist F such that $dF = Mdx + Ndy$.

Pfaff was one of Gauss' teachers at the beginning of the nineteenth century. He was one of the first mathematicians to pursue solutions to exact equations. One of the theorems he discovered is that almost any first order differential equation $Mdx + Ndy = 0$ can be multiplied by an integrating factor I to make the equation $IMdx + INdy = 0$ an exact equation. In other words, we can find I such that there exists F with $dF = IMdx + INdy$. I have not found a simple proof of this claim[8]

[8]this result of Pfaff's Theorem is a basic example of Frobenius Theorem, an important general PDE result.

Given the proposition above, it is clear we must seek an integrating factor I such that

$$\frac{\partial}{\partial y}\left[\ IM\ \right] = \frac{\partial}{\partial x}\left[\ IN\ \right].$$

Often for a particular problem we add some restriction to make the search for I less daunting. In each of the examples below I add a restriction on the search for I which helps us narrow the search[9]

Example 2.3.15. **Problem:** *find the solutions of*

$$\left(\ 3x + \frac{2}{y}\ \right)dx + \left(\ \frac{x^2}{y}\ \right)dy = 0 \qquad (\star)$$

by finding an integrating factor to make the equation exact.

Solution: *Since the problem only involves simple polynomials and rational functions the factor $I = x^Ay^B$ may suffice. Let us give it a try and see if we can choose a particular value for A, B to make I a proper integrating factor for the given problem. Multiply $\star$ by $I = x^Ay^B$,*

$$\left(\ 3x^{A+1}y^B + 2x^Ay^{B-1}\ \right)dx + \left(\ x^{A+2}y^{B-1}\ \right)dy = 0 \qquad (I\star)$$

Let $M = 3x^{A+1}y^B + 2x^Ay^{B-1}$ and $N = x^{A+2}y^{B-1}$. We need $\partial_y M = \partial_x N$, this yields:

$$3Bx^{A+1}y^{B-1} + 2(B-1)x^Ay^{B-2} = (A+2)x^{A+1}y^{B-1}$$

It follows that $3B = A+2$ and $2(B-1) = 0$. Thus $B = 1$ and $A = 1$. We propose $I = xy$ serves as an integrating factor for $\star$. Multiply by $\star$ by xy to obtain

$$\left(\ 3x^2y + 2x\ \right)dx + \left(\ x^3\ \right)dy = 0 \qquad (xy\star)$$

note that $F(x,y) = x^3y + x^2 = k$ has $\partial_x F = 3x^2y + 2x$ and $\partial_y F = x^3$ therefore $F(x,y) = x^3y + x^2 = k$ yield solutions to $xy\star$. These are also solutions for $\star$. However, we may have removed several solutions from the solution set when we multiplied by I. If $I = 0$ or if I is undefined for some points in the plane then we must consider those points separately and directly with $\star$. Note that $I = xy$ is zero for $x = 0$ or $y = 0$. Clearly $y = 0$ is not a solution for $\star$ since it is outside the domain of definition for $\star$ hence $y = 0$ is an **extraneous solution** *(IT'S NOT EVEN IN THE SOLUTION SET). On the other hand, $x = 0$ does solve $\star$ so we keep it in the general solution. Let us summarize: $\star$ has solutions of the form $x^3y + x^2 = k$ or $x \equiv 0$.*

Example 2.3.16. **Problem:** *find the solutions of $\frac{dy}{dx} + Py = Q$ by the method of exact equations. Assume that P, Q are differentiable functions of x.*

Solution: *in Pfaffian form this DEqn takes the form $dy + Pydx = Qdx$ or $(Py - Q)dx + dy = 0$. Generally, P, Q are not given such that this equation is exact. We seek an integrating factor I such that $I(Py - Q)dx + Idy = 0$ is exact. We need:*

$$\frac{\partial}{\partial y}\big[I(Py - Q)\big] = \frac{\partial}{\partial x}\big[I\big]$$

[9] It turns out that there are infinitely many integrating factors for a given inexact equation and we just need to find one that works. We find a few more helpful formulas towards the end of this section which show how to calculate an integrating factor in special cases.

Assume that I is not a function of y for the sake of discovery, and it follows that $IP = \frac{dI}{dx}$. This is solved by separation of variables: $\frac{dI}{I} = Pdx$ implies $\ln|I| = \int P\,dx$ yielding $I = exp(\int P\,dx)$. This means the integrating factor is an integrating factor. We gave several examples in the previous section.

The nice feature of the integrating factor $I = exp(\int P\,dx)$ is that when we multiply the linear differential equation $\frac{dy}{dx} + Py = Q$ we lose no solutions since $I \neq 0$. There are no extraneous solutions in this linear case.

There is deeper math to discover here. The problem of how to find a wise substitution is a fascinating topic that we could easily spend a semester developing better tools to solve such problems. In particular, if you wish to do further reading I recommend the text by Peter Hydon on symmetries and differential equations. Or, if you want a deeper discussion which is still primarily computational you might look at the text by Brian Cantwell. The basic idea is that if you know a *symmetry* of the differential equation it allows you to find special coordinates where the equation is easy to solve. Ignoring the symmetry part, this is what we did in this section, we found an integrating factor which transforms the given inexact equation to the simple exact equation $dF = 0$.

2.3.4 special integrating factors

Given $M(x,y)dx + N(x,y)dy = 0$ we say $\mu(x,y)$ is an

Definition 2.3.17. *Pffafian form of a differential equation*

Given $\omega = M(x,y)dx + N(x,y)dy$ we say $I(x,y)$ is an **integrating factor** for ω if $I\omega$ is an exact one-form. In other words, if $I(x,y)M(x,y)dx + I(x,y)N(x,y)dy = 0$ is an exact equation then I is an integrating factor for $Mdx + Ndy = 0$.

We say $\alpha = Pdx + Qdy$ is an **exact one-form** on $U \subseteq \mathbb{R}^2$ if there exists $F : U \to \mathbb{R}$ for which $dF = \partial_x F dx + \partial_y F dy = \alpha$. We can derive necessary conditions on the integrating factor. I will omit the (x,y) dependence in what follows for brevity. If

$$IMdx + INdy = 0$$

is an exact equation then we require the closed condition holds for IM and IN:

$$\frac{\partial}{\partial y}[IM] = \frac{\partial}{\partial x}[IN] \quad \Rightarrow \quad \frac{\partial I}{\partial y}M + I\frac{\partial M}{\partial y} = \frac{\partial I}{\partial x}N + I\frac{\partial N}{\partial x}$$

Consequently, the integrating factor I must solve:

$$\boxed{\frac{\partial I}{\partial y}M - \frac{\partial I}{\partial x}N = I\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)}.$$

It follows that we can solve for I provided certain conditions are met:

(1.) $I = I(x)$ then by assumption $\frac{\partial I}{\partial y} = 0$ and we face $-\frac{\partial I}{\partial x}N = I\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)$. However, since I only has x-dependence we find $\frac{\partial I}{\partial x} = \frac{dI}{dx}$ and hence

$$-\frac{dI}{dx}N = I\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) \quad \Rightarrow \quad \frac{dI}{I} = \frac{-1}{N}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)dx$$

Therefore,

$$\ln|I| = \int \frac{-1}{N}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx \quad \Rightarrow \quad \boxed{I = exp\left[\int \frac{-1}{N}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\right]}.$$

This choice of integrating factor requires that $\frac{\partial}{\partial y}\left[\frac{-1}{N}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\right] = 0$ since we required $I = I(x)$ as our initial assumption.

(2.) $I = I(y)$ then by assumption $\frac{\partial I}{\partial x} = 0$ and we face $\frac{\partial I}{\partial y}M = I\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)$. However, since I only has y-dependence we find $\frac{\partial I}{\partial y} = \frac{dI}{dy}$ and hence

$$\frac{dI}{dy}M = I\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) \quad \Rightarrow \quad \frac{dI}{I} = \frac{1}{M}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dy$$

Therefore,

$$\ln|I| = \int \frac{1}{M}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dy \quad \Rightarrow \quad \boxed{I = exp\left[\int \frac{1}{M}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dy\right]}.$$

This choice of integrating factor requires that $\frac{\partial}{\partial x}\left[\frac{1}{M}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\right] = 0$ since we need $I = I(y)$ as our starting assumption.

Example 2.3.18. Problem: *Solve* $(xy + y)dx + xdy = 0$

Solution: *identify* $M = xy + y$ *and* $N = x$ *thus*

$$\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 1 - (x+1) = -x \quad \Rightarrow \quad \frac{-1}{N}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) = \frac{-1}{x}(-x) = 1.$$

Following case **(1.)** *we propose* $I = exp(\int dx) = e^x$. *Multiplying* $(xy+y)dx + xdy = 0$ *by* e^x *yields:*

$$(xy + y)e^x dx + xe^x dy = 0 \quad \Rightarrow \quad d(yxe^x) = 0 \quad \Rightarrow \quad yxe^x = c \quad \Rightarrow \quad \boxed{y = \frac{c}{xe^x}}.$$

Example 2.3.19. Problem: *Solve* $(xy + x)dx + (y^2 + y)dy = 0$.

Solution: *identify* $M = xy + x$ *and* $N = y^2 + y$ *then*

$$\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 0 - x \quad \Rightarrow \quad \frac{1}{M}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) = \frac{1}{xy+x}(-x) = \frac{-1}{y+1}.$$

Therefore, following case **(2.)**, *we propose* $I = exp\left[\int\left(\frac{-1}{y+1}\right) dy\right] = exp\left[-\ln|y+1|\right] = \frac{1}{|y+1|}$. *As a general point of calculational ease, we can drop the absolute value bars and the resulting function will still serve as an integrating factor. Hence multiply* $(xy+x)dx + (y^2+y)dy = 0$ *by* $\frac{1}{y+1}$ *to obtain*

$$\left(\frac{xy + x}{y+1}\right) dx + \left(\frac{y^2+y}{y+1}\right) dy = 0 \quad \Rightarrow \quad xdx + ydy = 0 \quad \Rightarrow \quad d\left(\frac{1}{2}x^2 + \frac{1}{2}y^2\right) = 0$$

thus we find the solution $\boxed{x^2 + y^2 = c}$.

The theorem below is borrowed from Ritger and Rose page 53 of §2.5:

Theorem 2.3.20. *integrating factors are not unique*

If $u(x,y)$ is an integrating factor of $M\,dx + N\,dy = 0$ and if $dv = uM\,dx + uN\,dy$ then $u(x,y)F(v(x,y))$ is also an integrating factor for any continuous function F

To see how this is true, integrate F to obtain G such that $G'(v) = F(v)$. Observe $dG = G'(v)dv = F(v)dv$. However, we know $dv = uM\,dx + uN\,dy$ hence $dG = F(v)[uM\,dx + uN\,dy] = uFM\,dx + uFN\,dy$ which shows the DEqn $M\,dx + N\,dy = 0$ is made exact upon multiplication by uF. This makes uF and integrating factor as the theorem claims.

2.4 substitutions

In this section we discuss a few common substitutions. The idea of substitution is simply to transform a given problem to one we already know how to solve. Let me sketch the general idea before we get into examples: we are given

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

We propose a new dependent variable v which is defined by $y = h(x,v)$ for some function h. Observe, by the multivariate chain-rule,

$$\frac{dy}{dx} = \frac{d}{dx}h(x,v) = \frac{\partial h}{\partial x}\frac{dx}{dx} + \frac{\partial h}{\partial v}\frac{dv}{dx}$$

Hence, the substitution yields:

$$\frac{\partial h}{\partial x} + \frac{\partial h}{\partial v}\frac{dv}{dx} = f(x, h(x,v))$$

which, if we choose wisely, is simpler to solve.

Example 2.4.1. Problem: *solve* $\frac{dy}{dx} = (x+y-6)^2$. *(call this* $\star$*)*

Solution: *the substitution* $v = x+y-6$ *looks promising. We obtain* $y = v-x+6$ *hence* $\frac{dy}{dx} = \frac{dv}{dx} - 1$ *thus the DEQn* $\star$ *transforms to*

$$\frac{dv}{dx} - 1 = v^2 \quad \Rightarrow \quad \frac{dv}{dx} = v^2 + 1 \quad \Rightarrow \quad \frac{dv}{1+v^2} = dx \quad \Rightarrow \quad \tan^{-1}(v) = x + C$$

Hence, $\tan^{-1}(x+y-6) = x + C$ *is the general, implicit, solution to* $\star$*. In this case we can solve for* y *to find the explicit solution* $y = 6 + \tan(x+C) - x$*.*

Remark 2.4.2.

Generally the example above gives us hope that a DEqn of the form $\frac{dy}{dx} = F(ax+by+c)$ is solved through the substitution $v = ax + by + c$.

Example 2.4.3. Problem: *solve* $\frac{dy}{dx} = \frac{y/x+1}{y/x-1}$. *(call this* $\star$*)*

Solution: *the substitution* $v = y/x$ *looks promising. Note that* $y = xv$ *hence* $\frac{dy}{dx} = v + x\frac{dv}{dx}$ *by the product rule. We find* $\star$ *transforms to:*

$$v + x\frac{dv}{dx} = \frac{v+1}{v-1} \quad \Rightarrow \quad x\frac{dv}{dx} = \frac{v+1}{v-1} - v = \frac{v+1-v(v-1)}{v-1} = \frac{-v^2+2v+1}{v-1}$$

Hence, separating variables,

$$\frac{(v-1)\,dv}{-v^2+2v+1} = \frac{dx}{x} \quad \Rightarrow \quad -\frac{1}{2}\ln|v^2-2v-1| = \ln|x| + \tilde{C}$$

Thus, $\ln|v^2-2v-1| = \ln(1/x^2) + C$ *and after exponentiation and multiplication by* x^2 *we find the implicit solution* $\boxed{y^2 - 2xy - x^2 = K.}$

A differential equation of the form $\frac{dy}{dx} = F(y/x)$ is called **homogeneous**[10] . If we change coordinates by rescaling both x and y by the same scale then the ratio y/x remains invariant; $\bar{x} = \lambda x$ and $\bar{y} = \lambda y$ gives $\frac{\bar{y}}{\bar{x}} = \frac{\lambda y}{\lambda x} = \frac{y}{x}$. It turns out this is the reason the example above worked out so nicely, the coordinate $v = y/x$ is invariant under the rescaling symmetry.

Remark 2.4.4.

Generally the example above gives us hope that a DEqn of the form $\frac{dy}{dx} = F(y/x)$ is solved through the substitution $v = y/x$.

Example 2.4.5. Problem: *Solve* $y' + xy = xy^3$. *(call this* $\star$*)*

Solution: *multiply by* y^{-3} *to obtain* $y^{-3}y' + xy^{-2} = x$. *Let* $z = y^{-2}$ *and observe* $z' = -2y^{-3}y'$ *thus* $y^{-3}y' = \frac{-1}{2}z'$. *It follows that:*

$$\frac{-1}{2}\frac{dz}{dx} + xz = x \quad \Rightarrow \quad \frac{dz}{dx} - 2xz = -2x$$

Identify this is a linear ODE and calculate the integrating factor is e^{-x^2} *hence*

$$e^{-x^2}\frac{dz}{dx} - 2xe^{-x^2}z = -2xe^{-x^2} \quad \Rightarrow \quad d(e^{-x^2}z) = -2xe^{-x^2}dx$$

Conquently, $e^{-x^2}z = e^{-x^2} + C$ *which gives* $z = y^{-2} = 1 + Ce^{x^2}$. *Finally, solve for* y

$$\boxed{y = \frac{\pm 1}{\sqrt{1+Ce^{x^2}}}.}$$

Given an initial condition we would need to select either + or − as appropriate.

[10] this term is used several times in this course with differing meanings. The more common use arises in the discussion of linear differential equations.

Remark 2.4.6.

This type of differential equation actually has a name; a differential equation of the type $\frac{dy}{dx} + P(x)y = Q(x)y^n$ is called a **Bernoulli DEqn**. The procedure to solve such problems is as follows:

1. multiply $\frac{dy}{dx} + P(x)y = Q(x)y^n$ by y^{-n} to obtain $y^{-n}\frac{dy}{dx} + P(x)y^{-n+1} = Q(x)$,
2. make the substitution $z = y^{-n+1}$ and observe $z' = (1-n)y^{-n}y'$ hence $y^{-n}y' = \frac{1}{1-n}z'$,
3. solve the linear ODE in z; $\frac{1}{1-n}\frac{dz}{dx} + P(x)z = Q(x)$,
4. replace z with y^{-n+1} and solve if worthile for y.

Substutitions which change both the dependent and independent variable are naturally handled in the differential notation. If we replace $x = f(s,t)$ and $y = g(s,t)$ then $dx = f_s ds + f_t dt$ and $dy = g_s ds + g_t dt$. If we wish to transform $M(x,y)dx + N(x,y)dy$ into s,t coordinates we simply substitute the natural expressions:

$$M(x,y)dx + N(x,y)dy = M(f(s,t),g(s,t))\left[\frac{\partial f}{\partial s}ds + \frac{\partial f}{\partial t}dt\right] + N(f(s,t),g(s,t))\left[\frac{\partial g}{\partial s}ds + \frac{\partial g}{\partial t}dt\right].$$

Let us see how this works in a particular example:

Example 2.4.7. Problem: *solve $(x+y+2)dx + (x-y)dy = 0$. (call this $\star$)*

Solution: *the substitution $s = x+y+2$ and $t = x-y$ looks promising. Algebra yields $x = \frac{1}{2}(s+t-2)$ and $y = \frac{1}{2}(s-t-2)$ hence $dx = \frac{1}{2}(ds+dt)$ and $dy = \frac{1}{2}(ds-dt)$ thus $\star$ transforms to:*

$$s\frac{1}{2}(ds+dt) + t\frac{1}{2}(ds-dt) = 0 \quad \Rightarrow \quad (t+s)ds + (s-t)dt = 0 \quad \Rightarrow \quad \frac{dt}{ds} = \frac{t+s}{t-s}.$$

It follows, for $s \neq 0$,

$$\frac{dt}{ds} = \frac{t/s+1}{t/s-1}$$

Recall we solved this in Example 2.4.3 hence:

$$t^2 - 2st - s^2 = K \quad \Rightarrow \quad \boxed{(x-y)^2 - 2(x+y+2)(x-y) - (x+y+2)^2 = K.}$$

You can simplify that to $-2x^2 - 4xy - 8x + 2y^2 - 4 = K$. On the other hand, this DEqn is exact so it is considerably easier to see that $2x + \frac{x^2-y^2}{2} + xy = C$ is the solution. Multiply by -4 to obtain $-8x - 2x^2 + 2y^2 - 4xy = -4C$. It is the same solution as we just found through a much more laborious method. I include this example here to illustrate the method, naturally the exact equation approach is the better solution. Most of these problems do not admit the exact equation short-cut.

In retrospect, we were fortunate the transformed $\star$ was homogeneous. In Nagel Saff and Snider on pages 77-78 of the 5-th ed. a method for choosing s and t to insure homogeneity of the transformed DEqn is given.

Example 2.4.8. Problem: *solve* $\left[\frac{x}{\sqrt{x^2+y^2}} + y^2\right]dx + \left[\frac{y}{\sqrt{x^2+y^2}} - xy\right]dy = 0$. *(call this $\star$)*

Solution: *polar coordinates look promising here. Let $x = r\cos(\theta)$ and $y = r\sin(\theta)$,*

$$dx = \cos(\theta)dr - r\sin(\theta)d\theta, \qquad dy = \sin(\theta)dr + r\cos(\theta)d\theta$$

Furthermore, $r = \sqrt{x^2+y^2}$. We find $\star$ in polar coordinates,

$$\big[\cos(\theta) + r^2\sin^2(\theta)\big]\big[\cos(\theta)dr - r\sin(\theta)d\theta\big] + \big[\sin(\theta) - r^2\cos(\theta)\sin(\theta)\big]\big[\sin(\theta)dr + r\cos(\theta)d\theta\big] = 0$$

Multiply, collect terms, a few things cancel and we obtain:

$$dr + \big[-r^3\sin^3(\theta) - r^3\sin(\theta)\cos^2(\theta)\big]d\theta = 0$$

Hence,

$$dr - r^3\sin(\theta)d\theta = 0 \quad\Rightarrow\quad \frac{dr}{r^3} = \sin(\theta)d\theta \quad\Rightarrow\quad \frac{-1}{2r^2} = -\cos(\theta) + C.$$

Returning to Cartesian coordinates we find the implicit solution:

$$\boxed{\frac{1}{2(x^2+y^2)} = \frac{x}{\sqrt{x^2+y^2}} - C.}$$

Sometimes a second-order differential equation is easily reduced to a first-order problem. The examples below illustrate a technique called **reduction of order**.

Example 2.4.9. Problem: *solve $y'' + y' = e^{-x}$. (call this $\star$)*

Solution: *Let $y' = v$ and observe $y'' = v'$ hence $\star$ transforms to*

$$\frac{dv}{dx} - v = e^{-x}$$

multiply the DEqn above by the integrating factor e^x:

$$e^x\frac{dv}{dx} - ve^x = 1 \quad\Rightarrow\quad \frac{d}{dx}\Big[\ e^x v\ \Big] = 1$$

thus $e^x v = x + c_1$ and we find $v = xe^{-x} + c_1e^{-x}$. Then as $v = \frac{dy}{dx}$ we can integrate once more to find the solution:

$$y = \int \big[xe^{-x} + c_1e^{-x}\big]dx = -xe^{-x} - e^{-x} - c_1e^{-x} + c_2$$

cleaning it up a bit,

$$\boxed{y = -e^{-x}(x - 1 + c_1) + c_2.}$$

Remark 2.4.10.

Generally, given a differential equation of the form $y'' = F(y', x)$ we can solve it by a two-step process:

1. substitute $v = y'$ to obtain the first-order problem $v' = F(v, x)$. Solve for v.
2. recall $v = y'$, integrate to find y.

There will be two constants of integration. This is a typical feature of second-order ODE.

Example 2.4.11. Problem: *solve* $\frac{d^2y}{dt^2} + y = 0$. *(call this* $\star$*)*

Solution: *once more let* $v = \frac{dy}{dt}$. *Notice that*

$$\frac{d^2y}{dt^2} = \frac{dv}{dt} = \frac{dy}{dt}\frac{dv}{dy} = v\frac{dv}{dy}$$

thus $\star$ *transforms to the first-order problem:*

$$v\frac{dv}{dy} + y = 0 \quad \Rightarrow \quad vdv + ydy = 0 \quad \Rightarrow \quad \frac{1}{2}v^2 + \frac{1}{2}y^2 = \frac{1}{2}C^2.$$

assume the constant is positive we can express it as $C^2/2$, *note nothing is lost in doing this except the point solution* $y = 0, v = 0$. *Solving for* v *we obtain* $v = \pm\sqrt{C^2 - y^2}$. *However,* $v = \frac{dy}{dt}$ *so we find:*

$$\frac{dy}{\sqrt{C^2 - y^2}} = \pm dt \quad \Rightarrow \quad \sin^{-1}(y/C) = \pm t + \phi$$

Thus, $y = C\sin(\pm t + \phi)$. *We can just as well write* $y = A\sin(t + \phi)$. *Moreover, by trigonometry, this is the same as* $y = B\cos(t + \gamma)$, *it's just a matter of relabeling the constants in the general solution.*

Remark 2.4.12.

Generally, given a differential equation of the form $y'' = F(y)$ we can solve it by a two-step process:

1. substitute $v = y'$ and use the identity $\frac{dv}{dt} = v\frac{dv}{dy}$ to obtain the first-order problem $v\frac{dv}{dy} = F(y)$. Solve for v.
2. recall $v = y'$, integrate to find y.

There may be several cases possible as we solve for v, but in the end there will be two constants of integration.

I would like to also solve $\frac{d^2y}{dt^2} - y = 0$ and $\frac{d^2y}{dt^2} - 2a\frac{dy}{dt} + a^2y = 0$ where $a > 0$ via arguments similar to those given above. In fact, it would be nice to solve

$$ay'' + by' + cy = 0$$

for $a \neq 0$ and arbitrary $b, c \in \mathbb{R}$. We solve this later, but our method involves guessing in contrast to method above which I would call **derivation**. Deriving something from base principles is generally more difficult than simply finding a solution through some educated guessing. Much of what we do in this course falls under the general category of educated guessing.

2.5 physics and applications

I've broken this section into two parts. The initial subsection examines how we can use differential-equations techniques to better understand Newton's Laws and energy in classical mechanics. This sort of discussion is found in many of the older classic texts on differential equations. The second portion of this section is a collection of isolated application examples which are focused on a particular problems from a variety of fields.

2.5.1 physics

In physics we learn that $\vec{F}_{net} = m\vec{a}$ or, in terms of momentum $\vec{F}_{net} = \frac{d\vec{p}}{dt}$. We consider the one-dimensional problem hence we have no need of the vector notation and we generally are faced with the problem:

$$F_{net} = m\frac{dv}{dt} \qquad \text{or} \qquad F_{net} = \frac{dp}{dt}$$

where the momentum p for a body with mass m is given by $p = mv$ where v is the velocity as defined by $v = \frac{dx}{dt}$. The acceleration a is defined by $a = \frac{dv}{dt}$. It is also customary to use the dot and double dot notation for problems of classical mechanics. In particular: $v = \dot{x}$, $a = \dot{v} = \ddot{x}$. Generally the net-force can be a function of position, velocity and time; $F_{net} = F(x, v, t)$. For example,

1. the spring force is given by $F = -kx$
2. the force of gravity near the surface of the earth is given by $F = \pm mg$ ($\pm$ depends on interpretation of x)
3. force of gravity distance x from center of mass M given by $F = -\frac{GmM}{x^2}$
4. thrust force on a rocket depends on speed and rate at which mass is ejected
5. friction forces which depend on velocity $F = \pm bv^n$ ($\pm$ needed to insure friction force is opposite the direction of motion)
6. an external force, could be sinusoidal $F = A\cos(\omega t)$, ...

Suppose that the force only depends on x; $F = F(x)$ consider Newton's Second Law:

$$m\frac{dv}{dt} = F(x)$$

Notice that we can use the identity $\frac{dv}{dt} = \frac{dx}{dt}\frac{dv}{dx} = v\frac{dv}{dx}$ hence

$$mv\frac{dv}{dx} = F(x) \quad \Rightarrow \quad \int_{v_o}^{v_f} mv\,dv = \int_{x_o}^{x_f} F(x)\,dx \quad \Rightarrow \quad \boxed{\tfrac{1}{2}mv_f^2 - \tfrac{1}{2}mv_o^2 = \int_{x_o}^{x_f} F(x)\,dx.}$$

The equation boxed above is the **work-energy theorem**, it says the change in the kinetic energy $K = \frac{1}{2}mv^2$ is given by $\int_{x_o}^{x_f} F(x)\,dx$. which is the **work** done by the force F. This result holds for any net-force, however, in the case of a conservative force we have $F = -\frac{dU}{dx}$ for the **potential energy** function U hence the work done by F simplifies nicely

$$\int_{x_o}^{x_f} F(x)\,dx = -\int_{x_o}^{x_f} \frac{dU}{dx}\,dx = -U(x_f) + U(x_o)$$

and we obtain the **conservation of total mechanical energy** $\frac{1}{2}mv_f^2 - \frac{1}{2}mv_o^2 = -U(x_f) + U(x_o)$ which is better written in terms of energy $E(x, v) = \frac{1}{2}mv^2 + U(x)$ as $E(x_o, v_o) = E(x_f, v_f)$. The total energy of a conservative system is constant. We can also see this by a direct-argument on the differential equation below:

$$m\frac{dv}{dt} = -\frac{dU}{dx} \quad \Rightarrow \quad m\frac{dv}{dt} + \frac{dU}{dx} = 0$$

multiply by $\frac{dx}{dt}$ and use the identity $\frac{d}{dt}\left[\frac{1}{2}v^2\right] = v\frac{dv}{dt}$:

$$m\frac{dx}{dt}\frac{dv}{dt} + \frac{dx}{dt}\frac{dU}{dx} = 0 \quad \Rightarrow \quad \frac{d}{dt}\left[\frac{1}{2}mv^2\right] + \frac{dU}{dt} = 0 \quad \Rightarrow \quad \frac{d}{dt}\left[\frac{1}{2}mv^2 + U\right] = 0 \quad \Rightarrow \quad \boxed{\frac{dE}{dt} = 0.}$$

Once more we have derived that the energy is constant for a system with a net-force which is conservative. Note that as time evolves the expression $E(x, v) = \frac{1}{2}mv^2 + U(x)$ is invariant. It follows that the motion of the system is in described by an **energy-level** curve in the xv-plane. This plane is commonly called the **phase plane** in physics literature. Much information can be gleaned about the possible motions of a system by studying the energy level curves in the phase plane. We discuss this qualitative technique in Chapter 5.

We now turn to a mass m for which the net-force is of the form $F(x, v) = -\frac{dU}{dx} \mp b|v|^n$. Here we insist that $-$ is given for $v > 0$ whereas the $+$ is given for the case $v < 0$ since we assume $b > 0$ and this friction force ought to point opposite the direction of motion. Once more consider Newton's Second Law:

$$m\frac{dv}{dt} = -\frac{dU}{dx} \mp bv^n \quad \Rightarrow \quad m\frac{dv}{dt} + \frac{dU}{dx} = \mp b|v|^n$$

multiply by the velocity and use the identity as we did in the conservative case:

$$m\frac{dx}{dt}\frac{dv}{dt} + \frac{dx}{dt}\frac{dU}{dx} = \mp bv|v|^n \quad \Rightarrow \quad \frac{d}{dt}\left[\frac{1}{2}mv^2 + U\right] = \mp bv|v|^n \quad \Rightarrow \quad \boxed{\frac{dE}{dt} = \mp bv|v|^n.}$$

The friction force reduces the energy. For example, if $n = 1$ then we have $\frac{dE}{dt} = -bv^2$.

Remark 2.5.1.

The concept of energy is implicit within Example 2.4.11. I should also mention that the trick of multiplying by the velocity to reveal a conservation law is used again and again in the junior-level classical mechanics course.

2.5.2 applications

Example 2.5.2. Problem: *Suppose x is the position of a mass undergoing one-dimensional, constant acceleration motion. You are given that initially we have velocity v_o at position x_o and later we have velocity v_f at position x_f. Find how the initial and final velocities and positions are related.*

Solution: *recall that $a = \frac{dv}{dt}$ but, by the chain-rule we can write $a = \frac{dx}{dt}\frac{dv}{dx} = v\frac{dv}{dx}$. We are given that a is a constant. Separate variables, and integrate with respect to the given data*

$$a = \frac{dx}{dt}\frac{dv}{dx} = v\frac{dv}{dx} \quad \Rightarrow \quad a\,dx = v\,dv \quad \Rightarrow \quad \int_{x_o}^{x_f} a\,dx = \int_{v_o}^{v_f} v\,dv \quad \Rightarrow \quad a(x_f - x_o) = \frac{1}{2}\big(v_f^2 - v_o^2\big).$$

Therefore, $\boxed{v_f^2 = v_o^2 + 2a(x_f - x_o)}$. *I hope you recognize this equation from physics.*

Example 2.5.3. Problem: *suppose the population P grows at a rate which is directly proportional to the population. Let k be the proportionality constant. Find the population at time t in terms of the initial population P_o.*

Solution: *the given problem translates into the differential equation* $\frac{dP}{dt} = kP$ *with* $P(0) = P_o$. *Seperate variables and integrate, note* $P > 0$ *so I drop the absolute value bars in the integral,*

$$\frac{dP}{dt} = kP \quad \Rightarrow \quad \int \frac{dP}{P} = \int k\, dt \quad \Rightarrow \quad \ln(P(t)) = kt + C$$

Apply the initial condition; $\ln(P(0)) = k(0) + C$ *hence* $C = \ln(P_o)$. *Consequently* $\ln(P(t)) = \ln(P_o) + kt$. *Exponentiate to derive* $\boxed{P(t) = P_o e^{kt}}$.

In the example above I have in mind $k > 0$, but if we allow $k < 0$ that models exponential population decline. Or, if we think of P as the number of radioactive particles then the same mathematics for $k < 0$ models radioactive decay.

Example 2.5.4. Problem: *the voltage dropped across a resistor* R *is given by the product of* R *and the current* I *through* R. *The voltage dropped across a capacitor* C *depends on the charge* Q *according to* $C = Q/V$ *(this is actually the definition of capacitance). If we connect* R *and* C *end-to-end making a loop then they are in parallel hence share the same voltage:* $IR = \frac{Q}{C}$. *As time goes on the charge on* C *flows off the capacitor and through the resistor. It follows that* $I = -\frac{dQ}{dt}$. *If the capacitor initially has charge* Q_o *then find* $Q(t)$ *and* $I(t)$ *for the* **discharging capacitor**

Solution: *We must solve*

$$-R\frac{dQ}{dt} = \frac{Q}{C}$$

Separate variables, integrate, apply $Q(0) = Q_o$:

$$\frac{dQ}{Q} = -\frac{dt}{RC} \quad \Rightarrow \quad \ln|Q| = -\frac{t}{RC} + c_1 \quad \Rightarrow \quad Q(t) = \pm e^{c_1} e^{-t/RC} \quad \Rightarrow \quad \boxed{Q(t) = Q_o e^{-t/RC}}$$

Another application of first order differential equations is simply to search for curves with particular properties. The next example illustrates that concept.

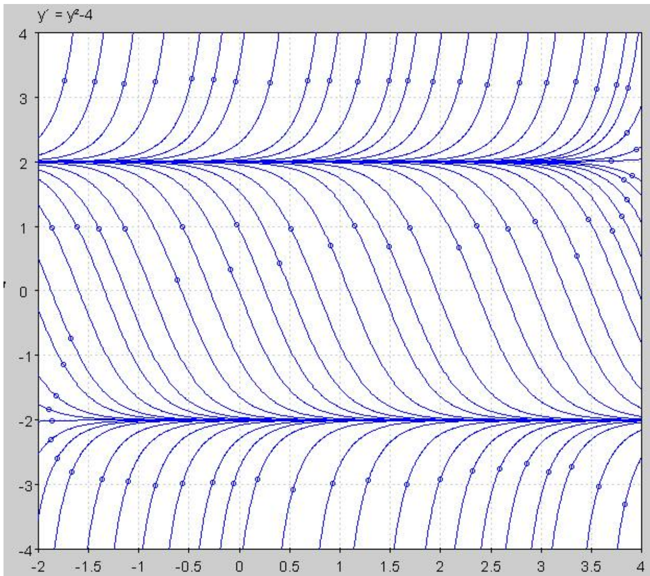
Example 2.5.5. Problem: *find a family of curves which are increasing whenever* $y < -2$ *or* $y > 2$ *and are decreasing whenever* $-2 < y < 2$.

Solution: *while many examples exist, the simplest example is one for which the derivative is quadratic in* y. *Think about the quadratic* $(y+2)(y-2)$. *This expression is positive for* $|y| > 2$ *and negative for* $|y| < 2$. *It follows that solutions to the differential equation* $\frac{dy}{dx} = (y+2)(y-2)$ *will have the desired properties. Note that* $y = \pm 2$ *are exceptional solutions for the give DEqn. Proceed* *by separation of variables, recall the technique of partial fractions,*

$$\begin{aligned}
\frac{dy}{(y+2)(y-2)} = dx \quad &\Rightarrow \quad \int \left[\frac{1}{4(y-2)} - \frac{1}{4(y+2)}\right] dy = \int dx \quad \star \\
&\Rightarrow \quad \ln|y-2| - \ln|y+2| = 4x + C \\
&\Rightarrow \quad \ln\left|\frac{y-2}{y+2}\right| = 4x + C \\
&\Rightarrow \quad \ln\left|\frac{y+2}{y+2} - \frac{4}{y+2}\right| = 4x + C \\
&\Rightarrow \quad \ln\left|1 - \frac{4}{y+2}\right| = 4x + C \\
&\Rightarrow \quad \left|1 - \frac{4}{y+2}\right| = e^{4x+C} = e^C e^{4x} \\
&\Rightarrow \quad 1 - \frac{4}{y+2} = \pm e^C e^{4x} = Ke^{4x} \\
&\Rightarrow \quad \frac{1}{y+2} = \frac{1-Ke^{4x}}{4} \\
&\Rightarrow \quad \boxed{y = -2 + \frac{4}{1-Ke^{4x}}, \quad \textit{for } K \neq 0.}
\end{aligned}$$

It is neat that $K = 0$ returns the exceptional solution $y = 2$ whereas the other exceptional solution is lost since we have division by $y + 2$ in the calculation above. If we had multiplied $\star$ by -1 then the tables would turn and we would recover $y = -2$ in the general formula.

The plot of the solutions below was prepared with pplane which is a feature of Matlab. To plot solutions to $\frac{dy}{dx} = f(x, y)$ you can put $x' = 1$ and $y' = f(x, y)$. This is an under-use of pplane. We discuss some of the deeper features towards the end of this chapter. Doubtless Mathematica will do these things, however, I don't have 10 hours to code it so, here it is:



If you study the solutions in the previous example you'll find that all solutions tend to either $y = 2$ or $y = -2$ in some limit. You can also show that all the solutions which cross the x-axis have inflection points at their x-intercept. We can derive that from the differential equation directly:

$$\frac{dy}{dx} = (y+2)(y-2) = y^2 - 4 \quad \Rightarrow \quad \frac{d^2y}{dx^2} = 2y\frac{dy}{dx} = 2y(y+2)(y-2).$$

We can easily reason when solutions have $y > 2$ or $-2 < y < 0$ they are concave up whereas solutions with $0 < y < 2$ or $y < -2$ are concave down. It follows that a solution crossing $y = 0, -2$ or 2 is at a point of inflection. Careful study of the solutions show that solutions do not cross $y = -2$ or $y = 2$ thus only $y = 0$ has solutions with genuine points of inflection.

Example 2.5.6. Problem: *suppose you are given a family S of curves which satisfy $\frac{dy}{dx} = f(x,y)$. Find a differential equation for a family of curves which are orthogonal to the given set of curves. In other words, find a differential equation whose solution consists of curves $S^{\perp}$ whose tangent vectors are perpendicular to the tangent vectors of curves in S at points of intersection.*

Solution: *Consider a point (x_o, y_o), note that the solution to $\frac{dy}{dx} = f(x,y)$ has slope $f(x_o, y_o)$ at that point. The perpendicular to the tangent has slope $-1/f(x_o, y_o)$. Thus, we should use the differential equation $\frac{dy}{dx} = -\frac{1}{f(x,y)}$ to obtain orthogonal trajectories.*

Let me give a concrete example of orthogonal trajectories:

Example 2.5.7. Problem: *find orthogonal trajectories of $xdx + ydy = 0$.*

Solution: *we find $\frac{dy}{dx} = \frac{-x}{y}$ hence the orthogonal trajectories are found in the solution set of $\frac{dy}{dx} = \frac{y}{x}$. Separate variables to obtain:*

$$\frac{dy}{y} = \frac{dx}{x} \quad \Rightarrow \quad \ln|y| = \ln|x| + C \quad \Rightarrow \quad y = \pm e^C x.$$

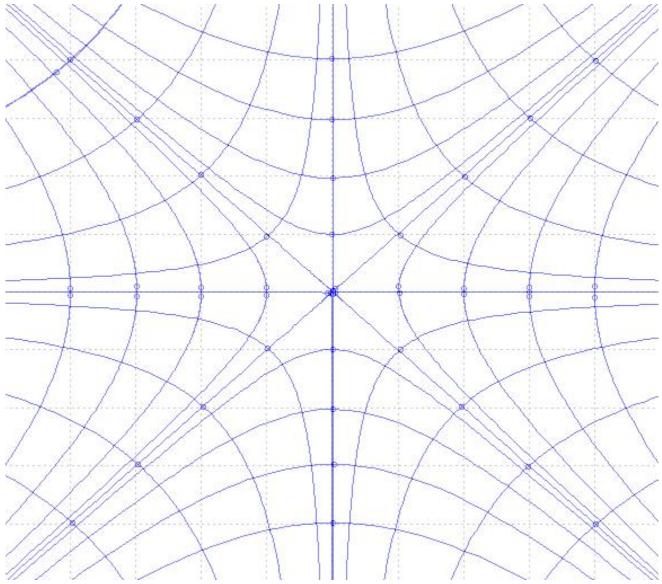
In other words, the orthogonal trajectories are lines through the origin $y = kx$. Technically, by our derivation, we ought not allow $k = 0$ but when you understand the solutions of $xdx + ydy = 0$ are simply circles $x^2 + y^2 = R^2$ it is clear that $y = 0$ is indeed an orthogonal trajectory.

Example 2.5.8. Problem: *find orthogonal trajectories of $x^2 - y^2 = 1$.*

Solution: *observe that the hyperbola above is a solution of the differential equation $2x - 2y\frac{dy}{dx} = 0$ hence $\frac{dy}{dx} = \frac{x}{y}$. Orthogonal trajectories are found from $\frac{dy}{dx} = \frac{-y}{x}$. Separate variables,*

$$\frac{dy}{y} = \frac{-dx}{x} \quad \Rightarrow \quad \ln|y| = -\ln|x| + C \quad \Rightarrow \quad y = k/x.$$

Once more, the case $k = 0$ is exceptional, but it is clear that $y = 0$ is an orthogonal trajectory of the given hyperbola.



Orthogonal trajectories are important to the theory of electrostatics. The field lines which are *integral curves* of the electric field form orthogonal trajectories to the *equpotential* curves. Or, in the study of heatflow, the isothermal curves are orthgonal to the curves which line-up with the flow of heat.

Example 2.5.9. Problem: *Suppose the force of friction on a speeding car is given by $F_f = -bv^2$. If the car has mass m and initial speed v_o and position x_o then find the velocity and position as a function of t as the car glides to a stop. Assume that the net-force is the friction force since the normal force and gravity cancel.*

Solution: *by Newton's second law we have $m\frac{dv}{dt} = -bv^2$. Separate variables, integrate. apply initital condition,*

$$\frac{dv}{v^2} = -\frac{bdt}{m} \quad \Rightarrow \quad \frac{-1}{v} = \frac{-bt}{m} + c_1 \quad \Rightarrow \quad \frac{-1}{v_o} = \frac{-b(0)}{m} + c_1 \quad \Rightarrow \quad c_1 = \frac{-1}{v_o}$$

Thus,

$$\frac{1}{v(t)} = \frac{bt}{m} + \frac{1}{v_o} \quad \Rightarrow \quad v(t) = \frac{1}{\frac{bt}{m} + \frac{1}{v_o}} \quad \Rightarrow \quad \boxed{v(t) = \frac{v_o}{\frac{btv_o}{m} + 1}}.$$

Since $v = \frac{dx}{dt}$ we can integrate the velocity to find the position

$$x(t) = c_1 + \frac{m}{b}\ln\left|1 + \frac{bv_ot}{m}\right| \quad \Rightarrow \quad x(0) = c_1 + \ln(1) = x_o \quad \Rightarrow \quad \boxed{x(t) = x_o + \frac{m}{b}\ln\left|1 + \frac{bv_ot}{m}\right|}.$$

Notice the slightly counter-intuitive nature of this solution, the position is unbounded even though the velocity tends to zero. Common sense might tell you that if the car slows to zero for large time then the total distance covered must be finite. Well, common sense fails, math wins. The point is that the velocity actually goes too zero too slowly to give bounded motion.

Example 2.5.10. Problem: *Newton's Law of Cooling states that the change in temperature T for an object is proportional to the difference between the ambient temperature R and T; in particular: $\frac{dT}{dt} = -k(T-R)$ for some constant k and R is the room-temperature. Suppose that $T(0) = 150$ and $T(1) = 120$ if $R = 70$, find $T(t)$*

Solution: *To begin let us examine the differential equation for arbitrary k and R,*

$$\frac{dT}{dt} = -k(T-R) \quad \Rightarrow \quad \frac{dT}{dt} + kT = kR$$

Identify that $p = k$ hence $I = e^{kt}$ and we find

$$e^{kt}\frac{dT}{dt} + ke^{kt}T = ke^{kt}R \quad \Rightarrow \quad \frac{d}{dt}\big[\, e^{kt}T\,\big] = ke^{kt}R \quad \Rightarrow \quad e^{kt}T = Re^{kt} + C \quad \Rightarrow \quad \boxed{T(t) = R + Ce^{-kt}.}$$

Now we may apply the given data to find both C and k, we already know $R = 70$ from the problem statement;

$$T(0) = 70 + C = 150 \qquad \& \qquad T(1) = 70 + Ce^{-k} = 120$$

Hence $C = 80$ which implies $e^{-k} = 5/8$ thus $e^k = 8/5$ and $k = \ln(8/5)$. Therefore, $\boxed{T(t) = 70 + 80e^{t\ln(5/8)}}$. *To understand this solution note that $\ln(5/8) < 0$ hence the term $80e^{t\ln(5/8)} \to 0$ as $t \to \infty$ hence $T(t) \to 70$ as $t \to \infty$. After a long time, Newton's Law of Cooling predicts objects will assume room temperature.*

Example 2.5.11. *Suppose you decide to have coffee with a friend and you both get your coffee ten minutes before the end of a serious presentation by your petty boss who will be offended if you start drinking during his fascinating talk on maximal efficiencies for production of widgets. You both desire to drink your coffee with the same amount of cream and you both like the coffee as hot as possible. Your friend puts the creamer in immediately and waits quitely for the talk to end. You on the other hand think you wait to put the cream in at the end of talk. Who has hotter coffee and why?* **Discuss.**

Example 2.5.12. Problem: *the voltage dropped across a resistor R is given by the product of R and the current I through R. The voltage dropped across an inductor L depends on the change in the current according to $L\frac{dI}{dt}$. An inductor resists a change in current whereas a resistor just resists current. If we connect R and L in series with a voltage source $\mathcal{E}$ then the Kirchoff's voltage law yields the differential equation*

$$\mathcal{E} - IR - L\frac{dI}{dt} = 0$$

Given that $I(0) = I_o$ find $I(t)$ for the circuit.

Solution: *Identify that this is a linear DE with independent variable t,*

$$\frac{dI}{dt} + \frac{R}{L}I = \frac{\mathcal{E}}{L}$$

The integrating factor is simply $\mu = e^{\frac{Rt}{L}}$ (using I here would be a poor notation). Multiplying the DEqn above by μ to obtain,

$$e^{\frac{Rt}{L}}\frac{dI}{dt} + \frac{R}{L}e^{\frac{Rt}{L}}I = \frac{\mathcal{E}}{L}e^{\frac{Rt}{L}} \quad \Rightarrow \quad \frac{d}{dt}\big[\, e^{\frac{Rt}{L}}I\,\big] = \frac{\mathcal{E}}{L}e^{\frac{Rt}{L}}$$

Introduce a dummy variable of integration τ *and integrate from* $\tau = 0$ *to* $\tau = t$,

$$\int_0^t \frac{d}{d\tau}\big[\, e^{\frac{R\tau}{L}} I\,\big] d\tau = \int_0^t \frac{\mathcal{E}}{L} e^{\frac{R\tau}{L}} d\tau \quad \Rightarrow \quad e^{\frac{Rt}{L}} I(t) - I_o = \int_0^t \frac{\mathcal{E}}{L} e^{\frac{R\tau}{L}} d\tau.$$

Therefore, $I(t) = I_o e^{\frac{-Rt}{L}} + e^{\frac{-Rt}{L}} \int_0^t \frac{\mathcal{E}}{L} e^{\frac{R\tau}{L}} d\tau$*. If the voltage source is constant then* $\mathcal{E}(t) = \mathcal{E}_o$ *for all* t *and the solution yields to* $I(t) = I_o e^{\frac{-Rt}{L}} + e^{\frac{-Rt}{L}} \frac{\mathcal{E}_o}{L} \frac{L}{R}\big(e^{\frac{Rt}{L}} - 1\big)$ *which simplifies to*

$$\boxed{I(t) = \left[I_o - \frac{\mathcal{E}_o}{R}\right] e^{\frac{-Rt}{L}} + \frac{\mathcal{E}_o}{R}}.$$

The **steady-state** *current found from letting* $t \to \infty$ *where we find* $I(t) \to \frac{\mathcal{E}_o}{R}$*. After a long time it is approximately correct to say the inductor is just a short-circuit. What happens is that as the current changes in the inductor a magnetic field is built up. The magnetic field contains energy and the maximum energy that can be stored in the field is governed by the voltage source. So, after a while, the field is approximately maximal and all the voltage is dropped across the resistor. You could think of it like saving money in a piggy-bank which cannot fit more than* $\mathcal{E}_o$ *dollars. If every week you get an allowance then eventually you have no choice but to spend the money if the piggy-bank is full and there is no other way to save.*

Example 2.5.13. Problem: *Suppose a tank of salty water has* $15kg$ *of salt disolved in* $1000L$ *of water at time* $t = 0$*. Furthermore, assume pure water enters the tank at a rate of* $10L/min$ *and salty water drains out at a rate of* $10L/min$*. If* $y(t)$ *is the number of* kg *of salt at time* t *then find* $y(t)$ *for* $t > 0$*. Also, how much salt is left in the tank when* $t = 20$ *(minutes). We suppose that this tank is arranged such that the concentration of salt is constant throughout the liquid in this* **mixing tank.**

Solution: *Generally when we work such a problem we are interested in the rate of salt enetering the tank and the rate of salt exiting the tank ;* $\frac{dy}{dt} = R_{in} - R_{out}$*. However, this problem only has a nonzero out-rate:* $R_{out} = \frac{10\,L}{min}\frac{y}{1000L} = \frac{y}{100min}$*. We omit the "min" in the math below as we assume* t *is in minutes,*

$$\frac{dy}{dt} = -\frac{y}{100} \quad \Rightarrow \quad \frac{dy}{y} = -\frac{dt}{100} \quad \Rightarrow \quad \ln|y| = -\frac{t}{100} + C \quad \Rightarrow \quad y(t) = ke^{-\frac{t}{100}}.$$

However, we are given that $y(0) = 15$ *hence* $k = 15$ *and we find*[11]*:*

$$\boxed{y(t) = 15e^{-0.01t}.}$$

Evaluating at $t = 20min$ *yields* $y(20) = 12.28\,kg$*.*

[11] to be physically explicit, $y(t) = (15kg)exp(\frac{-0.01t}{min})$, but the units clutter the math here so we omit them

Example 2.5.14. Problem: *Suppose a water tank has $100L$ of pure water at time $t = 0$. Suppose salty water with a concentration of $1.5kg$ of salt per L enters the tank at a rate of $8L/min$ and gets quickly mixed with the initially pure water. There is a drain in the tank where water drains out at a rate of $6L/min$. If $y(t)$ is the number of kg of salt at time t then find $y(t)$ for $t > 0$. If the water tank has a maximum capacity of $1000L$ then what are the physically reasonable values for the solution? For what t does your solution cease to be reasonable?*

Solution: *Generally when we work such a problem we are interested in the rate of salt enetering the tank and the rate of salt exiting the tank ; $\frac{dy}{dt} = R_{in} - R_{out}$. The input-rate is constant and is easily found from multiplying the given concentration by the flow-rate:*

$$R_{in} = \frac{1.5\,kg}{L}\frac{8\,L}{min} = \frac{12\,kg}{min}$$

notice how the units help us verify we are setting-up the model wisely. That said, I omit them in what follows to reduce clutter for the math. The output-rate is given by the product of the flow-rate $6L/min$ and the salt-concentration $y(t)/V(t)$ where $V(t)$ is the volume of water in L at time t. Notice that the $V(t)$ is given by $V(t) = 100 + 2t$ for the given flow-rates, each minute the volume increases by $2L$. We find (in units of kg and min):

$$R_{out} = \frac{6y}{100 + 2t}$$

Therefore, we must solve:

$$\frac{dy}{dt} = 12 - \frac{6y}{100 + 2t} \quad \Rightarrow \quad \frac{dy}{dt} + \frac{3}{50 + t}y = 12.$$

This is a linear ODE, we can solve it by the integrating factor method.

$$I(t) = exp\left(\int \frac{3dt}{50 + t}\right) = exp\Big(3\ln(50 + t)\Big) = (50 + t)^3.$$

Multiplying by I yields:

$$(50 + t)^3\frac{dy}{dt} + 3(50 + t)^2 y = 12(50 + t)^3 \quad \Rightarrow \quad \frac{d}{dt}\Big[(50 + t)^3 y\Big] = 12(50 + t)^3$$

Integrating yields $(50 + t)^3 y(t) = 3(50 + t)^4 + C$ hence $y(t) = 3(50 + t) + \frac{C}{(50+t)^3}$. The water is initially pure thus $y(0) = 0$ thus $0 = 150 + C/50^3$ which gives $C = -150(50)^3$. The solution is[12]

$$y(t) = 3(50 + t) - 150\left(\frac{50}{50 + t}\right)^3$$

Observe that $V(t) \leq 1000\,L$ thus we need $100 + 2t \leq 1000$ which gives $t \leq 450$. The solution is only appropriate physically for $0 \leq t \leq 450$.

[12] following the formatting of Example 7 of § 2.7 of Rice & Strange's Ordinary Differential Equations with Applications

Example 2.5.15. Problem: *suppose the population P grows at a rate which is directly proportional to the population. Let k_1 be the proportionality constant for the growth rate. Suppose further that as the population grows the death-rate is proportional to the square of the population. Suppose k_2 is the proportionality constant for the death-rate. Find the population at time t in terms of the initial population P_o.*

Solution: *the given problem translates into the IVP of $\frac{dP}{dt} = k_1P - k_2P^2$ with $P(0) = P_o$. Observe that $k_1P - k_2P^2 = k_1P(1 - k_2P/k_1)$. Introduce $C = k_1/k_2$. Separate variables:*

$$\frac{dP}{P(1-P/C)} = k_1 dt$$

Recall the technique of partial fractions:

$$\frac{1}{P(1-P/C)} = \frac{-C}{P(P-C)} = \frac{A}{P} + \frac{B}{P-C} \quad \Rightarrow \quad -C = A(P-C) + BP$$

Set $P = 0$ to obtain $-C = -AC$ hence $A = 1$ and set $P = C$ to obtain $-C = BC$ hence $B = -1$ and we find:

$$\int \left[\frac{1}{P} - \frac{1}{P-C}\right] dP = \int k_1 dt \quad \Rightarrow \quad \ln|P| - \ln|P-C| = k_1t + c_1$$

It follows that letting $c_2 = e^{c_1}$ and $c_3 = \pm c_2$

$$\left|\frac{P}{P-C}\right| = c_2 e^{k_1 t} \quad \Rightarrow \quad P = (P-C)c_3 e^{k_1 t}$$

hence, $P[1 - c_3e^{k_1t}] = -c_3Ce^{k_1t}$

$$P(t) = \frac{c_3Ce^{k_1t}}{c_3e^{k_1t} - 1} \quad \Rightarrow \quad P(t) = \frac{C}{1 - c_4e^{-k_1t}}$$

where I let $c_4 = 1/c_3$ for convenience. Let us work on writing this general solution in-terms of the initial population $P(0) = P_o$:

$$P_o = \frac{C}{1-c_4} \quad \Rightarrow \quad P_o(1-c_4) = C \quad \Rightarrow \quad P_o - C = P_oc_4 \quad \Rightarrow \quad c_4 = \frac{P_o - C}{P_o}.$$

This yields,

$$P(t) = \frac{C}{1 - \frac{P_o - C}{P_o}e^{-k_1t}} \quad \Rightarrow \quad \boxed{P(t) = C\left[\frac{P_o}{P_o - [P_o - C]e^{-k_1t}}\right]}$$

The quantity C is called the **carrying capacity** *for the system. As we defined it here it is given by the quotient of the birth-rate and death-rate constants $C = k_1/k_2$. Notice that as $t \to \infty$ we find $P(t) \to C$. If $P_o > C$ then the population decreases towards C whereas if $P_o < C$ then the population increases towards C. If $P_o = C$ then we have a special solution where $\frac{dP}{dt} = 0$ for all t, the* **equilbrium solution**. *A a bit of fun trivia, these models are notoriously incorrect for human populations. For example, in 1920 a paper by R. Pearl and L. J. Reed found $P(t) = \frac{210}{1+51.5e^{-0.03t}}$. The time t is the number of years past 1790 ($t = 60$ for 1850 for example). As discussed in Ritger and Rose page 85 this formula does quite well for 1950 where is well-approximates the population as 151 million. However, the carrying capacity of 210 million people is not even close to correct.*

Why? Because there are many factors which influence population which are simply not known. The same problem exists for economic models. You can't model game-changing events such as an interfering government. It doesn't flow from logic or optimal principles, political convenience whether it benefits or hurts a given market cannot be factored in over a long-term. Natural disasters also spoil our efforts to model populations and markets. That said, the exponential and logarthmic population models are important to a wide-swath of reasonably isolated populations which are free of chaotic events.

Example 2.5.16. Problem: *Suppose a raindrop falls through a cloud and gathers water from the cloud as it drops towards the ground. Suppose the mass of the raindrop is m and the suppose the rate at which the mass increases is proportional to the mass; $\frac{dm}{dt} = km$ for some constant $k > 0$. Find the equation of the velocity for the drop.*

Solution: *Newton's equation is $-mg = \frac{dp}{dt}$. This follows from the assumption that, on average, there is no net-momentum of the water vapor which adheres to the raindrop thus the momentum change is all from the gravitational force. Since $p = mv$ the product rule gives:*

$$-mg = \frac{dm}{dt}v + m\frac{dv}{dt} \quad \Rightarrow \quad -mg = kmv + m\frac{dv}{dt}$$

Consequently, dividing by m and applying the integrating factor method gives:

$$\frac{dv}{dt} + kv = -g \quad \Rightarrow \quad e^{kt}\frac{dv}{dt} + ke^{kt}v = -ge^{kt} \quad \Rightarrow \quad \frac{d}{dt}\left[e^{kt}v\right] = -ge^{kt}$$

Integrate to obtain $e^{kt}v = \frac{-g}{k}e^{kt} + C$ from which it follows $\boxed{v(t) = -\frac{g}{k} + Ce^{-kt}.}$ Consider the limit $t \to \infty$, we find $v_\infty(t) = -\frac{g}{k}$. This is called the **terminal velocity**. *Physically this is a very natural result; the velocity is constant when the forces balance. There are two forces at work here (1.) gravity $-mg$ and (2.) water friction $-kmv$ and we look at*

$$m\frac{dv}{dt} = -mg - kmv$$

If $v = -\frac{g}{k}$ then you obtain $ma = 0$. You might question if we should call the term $-kmv$ a "force". Is it really a force? In any event, you might note we can find the terminal velocity without solving the DEqn, we just have to look for an equilbrium of the forces.

Example 2.5.17. Problem: *Rocket flight. Rockets fly by ejecting mass with momentum to form thrust. We analyze the upward motion of a vertically launched rocket in this example. In this case Netwon's Second Law takes the form:*

$$\frac{d}{dt}\Big[\ mv\ \Big] = F_{external} + F_{thrust}$$

the external force could include gravity as well as friction and the thrust arises from conservation of momentum. Suppose the rocket expells gas downward at speed u relative the rocket. Suppose that the rocket burns mass at a uniform rate $m(t) = m_o - \alpha t$ and find the resulting equation of motion. Assume air friction is neglible.

Solution: *If the rocket has velocity v then the expelled gas has velocity $v - u$ relative the ground's frame of reference. It follows that:*

$$F_{thrust} = (v-u)\frac{dm}{dt}$$

Since $F_{external} = -mg$ and $\frac{dm}{dt} = -\alpha$ we must solve

$$\frac{d}{dt}\Big[\ mv\ \Big] = -mg + (v-u)\frac{dm}{dt} \quad \Rightarrow \quad \frac{dm}{dt}v + m\frac{dv}{dt} = -mg + v\frac{dm}{dt} - u\frac{dm}{dt}$$

Thus,

$$m\frac{dv}{dt} = -u\frac{dm}{dt} - mg$$

Suppose, as was given, that $m(t) = m_o - \alpha t$ hence $\frac{dm}{dt} = -\alpha$

$$(m_o - \alpha t)\frac{dv}{dt} = \alpha u - (m_o - \alpha t)g \quad \Rightarrow \quad \frac{dv}{dt} = \frac{\alpha u}{m_o - \alpha t} - g$$

We can solve by integration: assume $v(0) = 0$ as is physically reasonable,

$$v(t) = -u\ln(m_o - \alpha t) + u\ln(m_o) - gt = -u\ln\left(1 - \frac{\alpha t}{m_o}\right) - gt.$$

The initial mass m_o consists of fuel and the rocket itself: $m_o = m_f + m_r$. This model is only physical for time t such that $m_r \leq m_f + m_r - \alpha t$ hence $0 \leq t \leq m_f/\alpha$. Once the fuel is finished the empty rocket completes the flight by projectile motion. You can integrate $v = dy/dt$ to find the equation of motion. In particular:

$$\begin{aligned} y(t) &= \int_0^t \big[-u\ln(m_o - \alpha\tau) + u\ln(m_o) - g\tau\big]d\tau \qquad (2.1) \\ &= \left(-\frac{u}{\alpha}\Big[(\alpha\tau - m_o)\ln(m_o - \alpha\tau) - \alpha\tau\Big] + u\tau\ln(m_o) - \frac{1}{2}g\tau^2\right)\Bigg|_0^t \\ &= -\frac{u}{\alpha}\Big[(\alpha t - m_o)\ln(m_o - \alpha t) - \alpha t\Big] + ut\ln(m_o) - \frac{1}{2}gt^2 - \frac{m_o u}{\alpha}\ln(m_o) \\ &= ut - \frac{1}{2}gt^2 + u\frac{m_o}{\alpha}\left(1 - \frac{\alpha t}{m_o}\right)\ln\left(1 - \frac{\alpha t}{m_o}\right) \end{aligned}$$

Suppose $-\int_0^{\frac{m_f}{\alpha}} u\ln\left(1 - \frac{\alpha t}{m_o}\right)dt = A$ then $y(t) = A - \frac{1}{2}g\left(t - \frac{m_f}{\alpha}\right)^2$ for $t > \frac{m_f}{\alpha}$ as the rocket freefalls back to earth having exhausted its fuel.

Technically, if the rocket flies more than a few miles vertically then we ought to use the variable force of gravity which correctly accounts for the weaking of the gravitational force with increasing altitude. Mostly this example is included to show how variable mass with momentum transfer is handled.

Other interesting applications include chemical reactions, radioactive decay, blood-flow, other population models, dozens if not hundreds of modifications of the physics examples we've considered, rumor propogation, etc... the math here is likely found in any discipline which uses math to quantiatively describe variables. I'll conclude this section with an interesting example I found in Edwards and Penny's *Elementary Differential Equations with Boundary Value Problems*, the 3rd Ed.

Example 2.5.18. Problem: *Suppose a flexiible rope of length 4ft has 3ft coiled on the edge of a balcony and 1ft hangs over the edge. If at $t = 0$ the rope begins to uncoil further then find the velocity of the rope as it falls. Also, how long does it take for the rope to fall completely off the balcony. Suppose that the force of friction is neglible.*

Solution: *let x be the length of rope hanging off and suppose $v = dx/dt$. It follows that $x(0) = 1$ and $v(0) = 0$. The force of gravity is mg, note that if λ is the mass per unit length of the rope then $m = \lambda x$, thus:*

$$\frac{d}{dt}\big[\, mv \,\big] = mg \quad\Rightarrow\quad \frac{d}{dt}\big[\, \lambda x v \,\big] = \lambda x g \quad\Rightarrow\quad \lambda\frac{dx}{dt}v + \lambda x\frac{dv}{dt} = \lambda x g$$

The mass-density λ cancels and since $v = \frac{dx}{dt}$ and $\frac{dv}{dt} = \frac{dx}{dt}\frac{dv}{dx} = v\frac{dv}{dx}$ we find:

$$v^2 + xv\frac{dv}{dx} = xg \quad\Rightarrow\quad \left(\frac{v^2}{x} - g\right)dx + v\,dv = 0$$

In your text, in the discussion of special integrating factors, it is indicated that when $\frac{\partial_y M - \partial_x N}{N} = A(x)$ then $I = \int A\,dx$ is an integrating factor. Observe $M(x,v) = \frac{v^2}{x} - g$ and $N(x,v) = v$ hence $\frac{\partial_y M - \partial_x N}{N} = 2/x$ hence we calculate $I = exp(\int \frac{2dx}{x}) = exp(2\ln|x|) = x^2$. Don't believe it? Well, believe this:

$$\left(xv^2 - gx^2\right)dx + x^2 v\,dv = 0 \quad\Rightarrow\quad \frac{1}{2}x^2v^2 - \frac{1}{3}gx^3 = C.$$

Apply the intiial condtions $x(0) = 1$ and $v(0) = 0$ gives $-\frac{1}{3}g = C$ thus $\frac{1}{2}x^2v^2 - \frac{1}{3}gx^3 = -\frac{1}{3}g$. and we solve for $v > 0$,

$$v = \sqrt{\frac{2g}{3}\left(\frac{x^3-1}{x^2}\right)}$$

However, $v = \frac{dx}{dt}$ consequently, separating and integrating:

$$T = \sqrt{\frac{3}{2g}}\int_1^4 \frac{x\,dx}{\sqrt{x^3-1}} \approxeq 2.5\sqrt{\frac{3}{2g}} = 0.541\ s$$

by Wolfram Alpha. See Section 1.7 Example 6 of Edwards and Penny's Elementary Differential Equations with Boundary Value Problems, the 3rd Ed. for another approach involving Simpson's rule with 100 steps.

2.6 visualizations, existence and uniqueness

Given a curve in the $\mathbb{R}^2$ we have two general methods to describe the curve:

(1.) $F(x, y) = k$ as a **level curve** (2.) $\vec{r}(t) = \langle x(t), y(t)\rangle$ as a **parametrized curve**

As an example, we can either write $x^2 + y^2 = 1$ or $x = \cos(t), y = \sin(t)$. The parametric view has the advantage of capturing the direction or orientation of the curve. We have studied solutions of $Mdx + Ndy = 0$ in terms of cartesian coordinates and naturally our solutions were level curves. We now turn to ask what conditions ought to hold for the parametrization of the solution to $Mdx + Ndy = 0$.

Given a differentiable function of two variables $F : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ we may subsitute a differentiable path $\vec{r} : I \subseteq \mathbb{R} \to D$ to form the composite function $F \circ \vec{r} : I \subseteq \mathbb{R} \to \mathbb{R}$. If we denote $\vec{r}(t) = \langle x(t), y(t)\rangle$ then the multivariate chain-rule says:

$$\frac{d}{dt}F(x(t), y(t)) = \frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt}.$$

Suppose we have a level curve C which is the solution set of $F(x, y) = k$ and suppose C is the solution of $Mdx + Ndy = 0$ (call this $(\star xy)$. It follows that the level-function F must have $\partial_x F = M$ and $\partial_y F = N$. Continuing, suppose a parametrization of C is given by the set of functions $x, y : I \subseteq \mathbb{R} \to \mathbb{R}$ where $F(x(t), y(t)) = k$ for all $t \in I$. Notice that when we differentiate k with respect to t we obtain zero hence, applying the general chain rule to our context,

$$\frac{\partial F}{\partial x}(x(t), y(t))\frac{dx}{dt} + \frac{\partial F}{\partial y}(x(t), y(t))\frac{dy}{dt} = 0$$

for any parametrization of C. But, $\partial_x F = M$ and $\partial_y F = N$ hence

$$M(x(t), y(t))\frac{dx}{dt} + N(x(t), y(t))\frac{dy}{dt} = 0 \qquad (\star t)$$

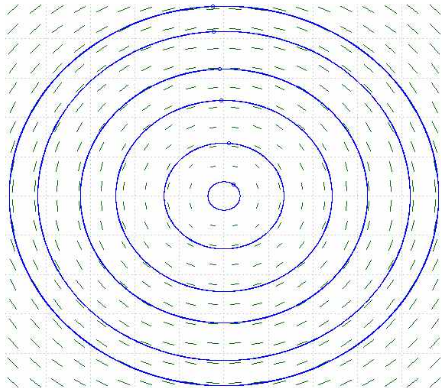
I probably cheated in class and just "divided by dt" to derive this from $Mdx + Ndy = 0$. However, that is just an abbreviation of the argument I present here. How should we solve $(\star t)$? Observe that the conditions

$$\frac{dx}{dt} = -N(x(t), y(t)) \qquad \& \qquad \frac{dy}{dt} = M(x(t), y(t))$$

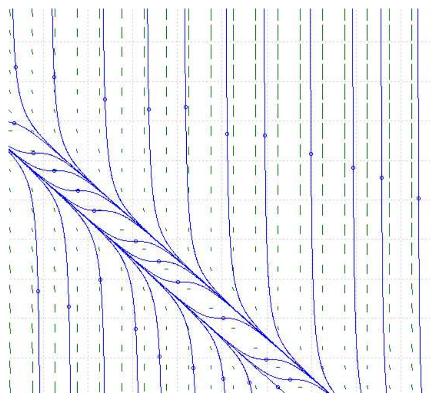
will suffice. Moreover, these conditions show that the solution of $Mdx + Ndy = 0$ is an streamline (or integral curve) of the vector field $\vec{G} = \langle -N, M\rangle$. Naturally, we see that $\vec{F} = \langle M, N\rangle$ is orthogonal to $\vec{G}$ as $\vec{F} \bullet \vec{G} = 0$. The solutions of $Mdx+Ndy = 0$ are perpendicular to the vector field $\langle M, N\rangle$.

There is an ambiguity we should face. Given $Mdx + Ndy = 0$ we can either view solutions as streamlines to the vector field $\langle -N, M\rangle$ or we could use $\langle N, -M\rangle$. The solutions of $Mdx+Ndy = 0$ do not have a natural direction unless we make some other convention or have some larger context. Therefore, as we seek to visualize the solutions of $Mdx + Ndy = 0$ we should either ignore the direction of the vector field $\langle -N, M\rangle$ or simply not plot the arrowheads. A plot of $\langle -N, M\rangle$ with directionless vectors is called an **isocline** plot for $Mdx + Ndy = 0$. Perhaps you looked at some isoclines in your second semester calculus course.

Example 2.6.1. Problem: *plot the isocline field for* $xdx + ydy = 0$ *and a few solutions.*
Solution: *use pplane with* $x' = -y$ *and* $y' = x$ *for the reasons we just derived in general.*

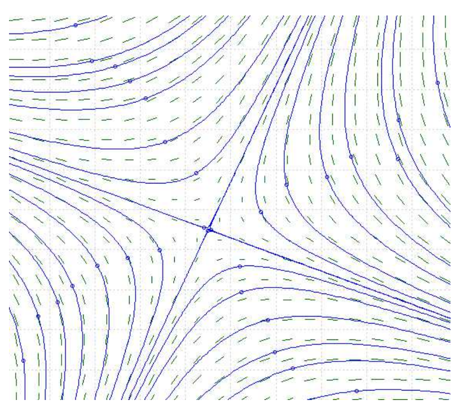


Example 2.6.2. Problem: *plot the isocline field for* $\frac{dy}{dx} = (x + y - 6)^2$ *and a few solutions.*
Solution: *in Pfaffian form we face* $(x + y - 6)^2 dx - dy$ *hence we use pplane with* $x' = -1$ *and* $y' = (x + y - 6)^2$.



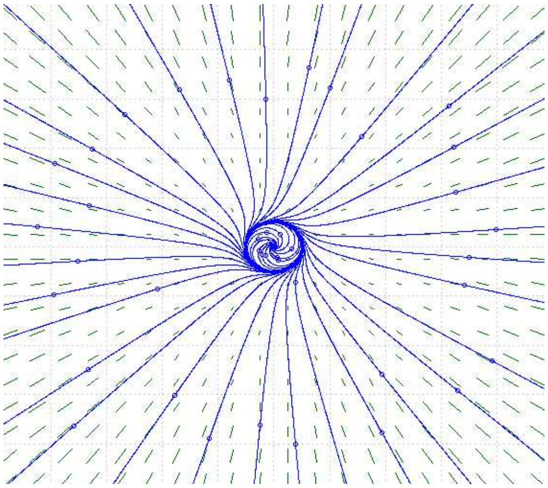
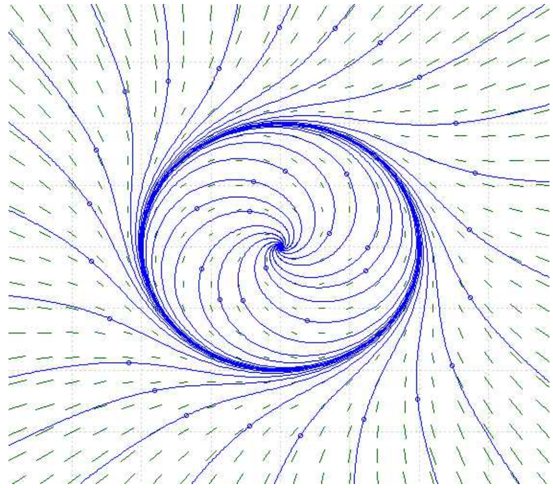
Recall that we found solutions $y = 6 + \tan(x + C) - x$ *in Example 2.4.1. This is the plot of that.*

Example 2.6.3. Problem: *plot the isocline field for* $(x + y + 2)dx + (x - y)dy = 0$ *and a few solutions.*
Solution: *in Pfaffian form we face* $(x + y - 6)^2 dx - dy$ *hence we use pplane with* $x' = -1$ *and* $y' = (x + y - 6)^2$.



Recall that we found solutions $2x + \frac{x^2 - y^2}{2} + xy = C$ *in Example 2.4.7. This is the plot of that.*

Example 2.6.4. **Problem:** *plot the isocline field for* $\frac{dy}{dx} = \frac{y^3+x^2y-y-x}{xy^2+x^3+y-x}$ *and a few solutions.*
Solution: *in Pfaffian form we face* $(y^3+x^2y-y-x)dx - (xy^2+x^3+y-x)dy$ *hence we use pplane with* $x' = xy^2+x^3+y-x$ *and* $y' = y^3+x^2y-y-x$.

See my handwritten notes (38-40) for the solution of this by algebraic methods. It is a beautiful example of how polar coordinate change naturally solves a first order ODE with a rotational symmetry. Also, notice that all solutions asymptotically are drawn to the unit circle. If the solution begins inside the circle it is drawn outwards to the circle whereas all solutions outside the circle spiral inward.

If you study the plots I just gave you will notice that at most points there is just one solution the flows through. However, at certain points there are multiple solutions that intersect. When there is just one solution at a given point (x_o, y_o) then we say that the solution is **unique**. It turns out there are simple theorems that capture when the solution is unique for a general first order ODE of the form $\frac{dy}{dx} = f(x,y)$. I will not prove these here[13]

Theorem 2.6.5. *existence of solution(s)*

> Suppose f is continuous on a rectangle $R \subset \mathbb{R}^2$ then **at least** one solution exists for $\frac{dy}{dx} = f(x,y)$ at each point in R. Moreover, these solutions exist on all of R in the sense that they reach the edge of R.

This is a dumbed-down version of the theorem given in the older texts like Rabenstein or Ritger & Rose. See pages 374-378 of Rabenstein or Chapter 4 of Ritger & Rose. You can read those if you wish to see the man behind the curtain here.

Theorem 2.6.6. *uniqueness of solution*

> Suppose f is continuous on a rectangle $R \subset \mathbb{R}^2$ and $\frac{\partial f}{\partial y}(x_o, y_o) \neq 0$ then **there exists a unique** solution near (x_o, y_o).

Uniqueness can be lost as we get too far away from the point where $\frac{\partial f}{\partial y}(x_o, y_o) \neq 0$. The solution is separated from other solutions near (x_o, y_o), but it may intersect other solutions as we travel away from the given point.

[13] see Rosenlicht's *Introduction to Analysis for a proof of this theorem, you need ideas from advanced calculus and real analysis to properly understand the proof.*

Example 2.6.7. *Solve $\frac{dy}{dx} = y^2$ and analyze how the uniqueness and existence theorems are validated. This nonlinear DEqn is easily solved by separation of variables: $dy/y^2 = dx$ hence $-1/y = x + C$ or $y = \frac{-1}{x+C}$. We also have the solution $y = 0$. Consider,*

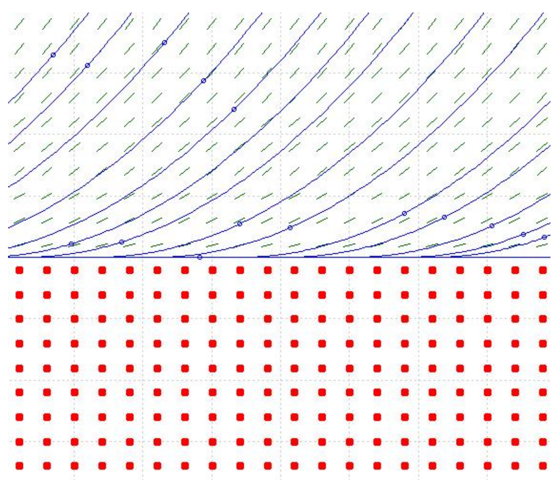
$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}[y^2] = 2y$$

Thus all points off the $y = 0$ (aka x-axis) should have locally unique solutions. In fact, it turns out that the solution $y = 0$ is also unique in this case. Notice that the theorem does not forbid this. The theorem on uniqueness only indicates that it is **possible** *for multiple solutions to exist at a point where $\frac{\partial f}{\partial y}(x_o, y_o) = 0$. It is important to not over-extend the theorem. On the other hand, the existence theorem says that solutions should extend to the edge of $\mathbb{R}^2$ and that is clearly accomplished by the solutions we found. You can think of $y = 0$ as reaching the horizontal infinities of the plane whereas the curves $y = \frac{-1}{x+C}$ have vertical asymptotes which naturally extend to the vertical infinities of the plane. (these comments are heuristic !)*

Example 2.6.8. *Solve $\frac{dy}{dx} = 2\sqrt{y}$ and analyze how the uniqueness and existence theorems are validated. Observe that $\frac{dy}{2\sqrt{y}} = dx$ hence $\sqrt{y} = x + C$ and we find $y = (x + C)^2$. Note that the solutions reach points on $\mathbb{R} \times [0, \infty)$ however the solutions do not have $y < 0$. The existence theorem suggests solutions should exist on $\mathbb{R} \times [0, \infty)$ and this is precisely what we found. On the other hand, for uniqueness, consider: $f(x, y) = 2\sqrt{y}$*

$$\frac{\partial f}{\partial y} = \frac{1}{\sqrt{y}}$$

We can expect unique solutions at points with $y \neq 0$, however, we **may** *find mutiple solutions at points with $y = 0$. Indeed, note that $y = 0$ is a solution and at any point $(a, 0)$ we also have the solution $y = (x - a)^2$. At each point along the x-axis we find two solutions intersect the point. Moreover, if you look at an open interval centered at $(a, 0)$ you'll find infinitely many solutions which flow off the special solution $y = 0$. Note, in the plot below, the pplane tool illustrates the points outside the domain of defintion by the red dots:*



Example 2.6.9. *Consider the first order linear ODE $\frac{dy}{dx} + P(x)y = Q(x)$. Identify that $\frac{dy}{dx} = Q(x) - P(x)y = f(x, y)$. Therefore, $\frac{\partial f}{\partial y} = -P(x)$. We might find there are multiple solutions for points with $P(x) = 0$. However, will we? Discuss.*

If a first order ODE does not have the form[14] $\frac{dy}{dx} + P(x)y = Q(x)$ then it is said to be **nonlinear**. Often the nonlinear ODEs we have studied have possessed unique solutions at most points. However, the unique solutions flow into some exceptional solution like $y = 0$ in Example 2.6.8 or the unit circle $x^2 + y^2 = 1$ or origin $(0, 0)$ in Example 2.6.4. These exceptional solutions for nonlinear problems are called **singular solutions** and a point like the origin in Example 2.6.4 is naturally called a **singular point**. That said, we will discuss a more precise idea of singular point for systems of ODEs later in this course. The study of nonlinear problems is a deep an interesting subject which we have only scratched the surface of here. I hope you see by now that the resource of pplane allows you to see things that would be very hard to see with more naive tools like a TI-83 or uncoded Mathematica.

2.7 practice problems

PP 1 (Separation of Variables) Solve the differential equations below. If possible, find the explicit solution, otherwise find an implicit general solution.

(a.) $\frac{dy}{dx} = (x+1)^2$

(b.) $dx - x^2 dy = 0$

(c.) $e^x \frac{dy}{dx} = 2x$

(d.) $\frac{dy}{dt} + 2xt = 0$

(e.) $\frac{dy}{dx} = \frac{y+1}{x}$

(f.) $\frac{dx}{dy} = \frac{1+2y^2}{y \sin x}$

(g.) $x^2 y^2 dy = (y+1)dx$

(h.) $\frac{dy}{dx} = \left(\frac{2y+3}{4x+5}\right)^2$

(i.) $\sec x dy = x \cot y dx$

(j.) $x\sqrt{1-z^2}dx = dz$

(k.) $\sec y \frac{dy}{dx} + \sin(x-y) = \sin(x+y)$

(l.) $\frac{dy}{dx} = \frac{xy+3x-y-3}{xy-2x+4y-8}$

(m.) $(e^x + e^{-x})\frac{dw}{dx} = w^2$

(n.) $(x+\sqrt{x})\frac{dy}{dx} = y + \sqrt{y}$.

[14] could also write $\frac{dy}{dx} = Q(x) - P(x)y$ or $dy = (Q(x) - P(x)y)dx$ etc... the key is that the expression has y and y' appearing linearly when the DEqn is written in $\frac{dy}{dx}$ notation

PP 2 (Initial Value Problems) Use separation of variables to solve the IVPs below:

(a.) $\dfrac{dy}{dt} + ty = y$ with $y(1) = 3$.

(b.) $\dfrac{dx}{dy} = 4(x^2+1)$ with $x(\pi/4) = 1$.

(c.) $\dfrac{dy}{dx} = \dfrac{y^2-1}{x^2-1}$ with $y(2) = 2$

(d.) $y' + 2y = 1$ with $y(0) = 5/2$.

PP 3 (substitution of form $u = ax + by + c$) Solve the following by making an appropriate substitution and using separation of variables,

(a.) $\dfrac{dy}{dx} = (x+y+1)^2$

(b.) $\dfrac{dy}{dx} = \tan^2(x+y)$

(c.) $\dfrac{dy}{dx} = \dfrac{1-x-y}{x+y}$

(d.) $\dfrac{dy}{dx} = 1 + e^{y-x+5}$

PP 4 (homogeneous equations, try $y = ux$ or $x = vy$ on $M(x,y)dx + N(x,y)dy = 0$ where M and N are homogeneous functions of same degree)

(a.) $(x-y)dx + xdy = 0$

(b.) $\dfrac{dy}{dx} = \dfrac{y-x}{y+x}$

(c.) $y\dfrac{dx}{dy} = x + 4ye^{-2x/y}$

(d.) $(x^2 + xy - y^2)dx + xydy = 0$

PP 5 (exact equations) If the DEqn below is exact then solve it, otherwise explain why the given DEqn is not exact.

(a.) $(2x+y)dx - (x+6y)dy = 0$

(b.) $(2xz^2 - 3)dx + (2zx^2 + 4)dz = 0$

(c.) $(\sin y - y\sin x)dx + (\cos x + x\cos y - y)dy = 0$

(d.) $\left(2y - \dfrac{1}{x} + \cos 3x\right)\dfrac{dy}{dx} + \dfrac{y}{x^2} - 4x^3 + 3y\sin 3x = 0$

(e.) $\left(1 + \ln x + \frac{y}{x}\right)dx = (1 - \ln x)dy$

(f.) $(\theta^3 + \beta^3)d\theta + 3\theta\beta^2 d\beta = 0$

(g.) $(3x^2y + e^y)dx + (x^3 + xe^y - 2y)dy = 0$

(h.) $(e^y + 2xy\cosh x)y' + xy^2\sinh x + y^2\cosh x = 0$

(i.) $(2y\sin x\cos x - y + 2y^2e^{xy^2})dx = (x - \sin^2 x - 4xye^{xy^2})dy$

(j.) $\left(\dfrac{1}{x} + \dfrac{1}{x^2} - \dfrac{t}{x^2+t^2}\right)dx + \left(te^t + \dfrac{x}{x^2+t^2}\right)dt = 0$

PP 6 (exact equations with integrating factor) A general form of an integrating factor is suggested. Find the specific form I which serves as an integrating factor and solve the DEqn $Mdx + Ndy = 0$ by solving the exact equation $IMdx + INdy = 0$)

(a.) $6xydx + (4y + 9x^2)dy = 0$ given $I = y^A$

(b.) $y(x + y + 1)dx + (x + 2y)dy = 0$ given $I = e^{Ax}$

(c.) $(x^2 + 2xy - y^2)dx + (y^2 + 2xy - x^2)dy = 0$ given $I = (x + y)^A$

(d.) $y(4xy^5 + 3)dx - x(2xy^5 + 7)dy = 0$ given $I = x^A y^B$

PP 7 (linear first order DEqn) Solve the linear first order ODEqn given below and state the interval on which the solution is defined. If given an initial value, then fit the given data to the explicit solution.

(a.) $\dfrac{dy}{dx} + y = e^{3x}$

(b.) $y' + 3x^2 y = x^2$

(c.) $\dfrac{dx}{dy} = x + y$

(d.) $x\dfrac{dy}{dx} + 2y = 3$

(e.) $xdy = x \sin x - y)dx$

(f.) $x\dfrac{dy}{dx} + 4y = x^3 - x$

(g.) $x^2 y' + x(x + 2)y = e^x$

(h.) $\cos^2 x \sin x dy + (y \cos^3 x - 1)dx = 0$

(i.) $\dfrac{dr}{d\theta} + r \sec\theta = \cos\theta$

(j.) $L\dfrac{di}{dt} + Ri = E$ where L, R, E are nonzero constants and $i(0) = i_o$

PP 8 (Bernoulli's Equation). If the DEqn has form $\frac{dy}{dx} + P(x)y = f(x)y^n$ for some real n then it is called a Bernoulli Equation. These can be solved by a $w = y^{1-n}$ substitution, we assume $n \neq 0, 1$. Solve the following:

(a.) $\dfrac{dy}{dx} - y = e^x y^2$

(b.) $x\dfrac{dy}{dx} - (1 + x)y = xy^2$

(c.) $3(1 + x^2)\dfrac{dy}{dx} = 2xy(y^3 - 1)$

PP 9 (Ricatti's Equation). If $\frac{dy}{dx} = P(x) + Q(x)y + R(x)y^2$ then the given DEqn is a Ricatti Equation. If y_1 is a known solution then the substitution $v = y_1 + u$ turns the problem into a Bernoulli Equation with $n = 2$. Given the Ricatti Equation below with known solution y_1, solve it. Or, if no y_1 is given then figure one out then solve it.

(a.) $\dfrac{dy}{dx} = 1 - x - y + xy^2$, $y_1 = 1$

(b.) $\frac{dy}{dx} = 2x^2 + y/x - 2y^2$, $y_1 = x$

(c.) $\frac{dy}{dx} = \sec^2 x - (\tan x)y + y^2$, $y_1 = \tan x$

(d.) $\frac{dy}{dx} = 9 + 6y + y^2$

PP 10 (Clauiraut Equation) Let f be a smooth function. The differential equation $y = xy' + f(y')$ is known as a Clairaut Equation. Show that $y = cx + f(c)$ serves as a solution to Clairaut Equation for any $c \in \mathbb{R}$. Furthermore, show

$$x = -f'(t), \qquad \& \qquad y = f(t) - tf'(t)$$

give a parametric solution to Clairaut Equation. If $f''(t) \neq 0$ then the parametric solution describes a solution not found in the linear family and as such it is known as the **singular solution**. Solve the Clairaut Equations below by finding both their linear solutions and the singular solution.

(a.) $y = xy' + (y')^{-2}$

(b.) $y = (x + 4)y' + (y')^2$

(c.) $y - xy' = \ln y'$

PP 11 Solve $2xy\frac{dy}{dx} + 2y^3 = 3x - 6$ by making a substitution of $v = y^2$.

PP 12 Solve $y'' = 2x(y')^2$ by making a $v = y'$ substitution.

PP 13 Solve $\frac{dy}{dx} = e^{x-y}\cosh x$

PP 14 Solve $\frac{dy}{dx} = \frac{y^2 + 4y + 5}{x^2 - 3x - 4}$

PP 15 Solve $(y + \sin^{-1}(x))dx + \left(x + \frac{1}{1 + y^2}\right) dy = 0$

PP 16 Solve $\frac{dy}{dx} = \frac{y^2 + 2xy}{x^2}$. Hint: write as $\frac{dy}{dx} = F(y/x)$.

PP 17 Find the explicit solution of $\frac{dy}{dx} = \frac{e^x}{y}$ for which $y(0) = -2$.

PP 18 Find the explicit general solution of $\frac{dy}{dx} + (\tan x)y = \sec x$.

PP 19 Find a continuous solution of $\frac{dy}{dx} = |x - 3|$ which contains the origin.

PP 20 Find a continuous solution of $\frac{dy}{dx} = \sqrt{(x-3)^2} - \sqrt{(x-4)^2}$ which contains the origin.

PP 21 Find the explicit solution of $\frac{dy}{dx} - \frac{3y}{x} = \sqrt{x}$.

PP 22 Note that $\mu = 1/(yx^2)$ is an integrating factor for the following differential equation:

$$(3x^2y+y^2)dx+(x^2y^2-xy)dy = 0,$$

Find the general solution to the differential equation. If there are any exceptional solutions be sure to point them out.

PP 23 Find the solution of $(2+ye^{xy})\,dx + (y+xe^{xy})\,dy = 0$ through (x_o, y_o). Hint: this is an exact DEqn thus you can use the theorem involving line integrals to build the desired solution

PP 24 Solve $(2y\sec^2(x^2y))dx + (x\sec^2(x^2y) + \frac{1}{x}e^y)dy = 0$.
Hint: $\mu = x$ is an integrating factor for this inexact equation.

PP 25 Solve $\frac{dy}{dx} = \frac{y^2+2xy}{x^2}$.
Hint: write as $\frac{dy}{dx} = F(y/x)$)

PP 26 Solve

$$\frac{dy}{dx} = y(xy^3-1).$$

PP 27 Solve the following by making a subsitution which replaces x and y with polar coordinates r and θ. Please give your answer in terms of r, θ.

$$[2x(x^2+y^2)+y]dx + [2y(x^2+y^2)-x]dy = 0.$$

PP 28 Find the implicit solution of:

$$\left(1+2xy^2-\frac{1}{x^2+4}\right)dx+\left(2y+2x^2y-\frac{1}{1-y^2}\right)dy = 0.$$

PP 29 Find the potential energy function for the conservative vector field:

$$\vec{F}(x,y) = \left\langle 1+2xy^2-\frac{1}{x^2+4}, 2y+2x^2y-\frac{1}{1-y^2}\right\rangle$$

PP 30 Consider the differential equation $Pdx + Qdy = 0$ where $P_y = Q_x$ on a simply connected region $U \subseteq \mathbb{R}^2$. Use Calculus III to prove there exists $F : U \to \mathbb{R}$ for which $dF = Pdx+Qdy$ and explain why $F(x,y) = C$ serves to solve $Pdx + Qdy = 0$.

PP 31 Consider the differential equation $Pdx + Qdy = 0$ where $P_y \neq Q_x$ on a simply connected region $U \subseteq \mathbb{R}^2$. Use Calculus III to prove there cannot exist $F : U \to \mathbb{R}$ for which $dF = Pdx + Qdy$.

PP 32 Consider $\omega = \dfrac{-ydx+xdy}{x^2+y^2}$.

(a.) Calculate $\int_C \omega$ where C is the CCW unit-circle.

(b.) We say ω is closed if $d\omega = 0$. Show ω is closed.

(c.) We say ω is exact if there exists F for which $dF = \omega$ on $U \subseteq \mathbb{R}^2$. Is ω exact on $\mathbb{R}^2$?

(d.) On which simply connected subsets of $\mathbb{R}^2$ is ω exact ?

PP 33 The wedge product follows the usual rules of algebra except it satisfies $dx \wedge dx = 0$ and $dx \wedge dy = -dy \wedge dx$ etc. Calculate $(a_1 dx + a_2 dy + a_3 dz) \wedge (b_1 dx + b_2 dy + b_3 dz)$ and comment on the meaning of the constants c_1, c_2, c_3 with

$$(a_1 dx + a_2 dy + a_3 dz) \wedge (b_1 dx + b_2 dy + b_3 dz) = c_1 dy \wedge dz + c_2 dz \wedge dx + c_3 dx \wedge dy$$

can you recognize $\langle c_1, c_2, c_3 \rangle$ as it relates to $\langle a_1, a_2, a_3 \rangle$ and $\langle b_1, b_2, b_3 \rangle$?

PP 34 The exterior derivative of a one-form $\omega = Adx + Bdy + Cdz$ is given by $d\omega = dA \wedge dx + dB \wedge dy + dC \wedge dz$ where $dA = (\partial_x A)dx + (\partial_y A)dy + (\partial_z A)dz$ is the total differential you know and love from Calculus III. Let f be a smooth function of x, y, z. Show that $d(df) = 0$. To which identity of vector calculus does your calculation correspond ?

PP 35 A differential equation $Mdx + Ndy = 0$ is exact if there exists F for which $dF = Mdx + Ndy$. Since $d(dF) = 0$ is an identity of the exterior calculus we can check on the exactness of a given differential equation in Pfaffian form by taking its exterior derivative. Determine if the differential equations below are exact by taking the exterior derivative of the differential equation:

(a.) $ydx + x^2 dy = 0$

(b.) $y \sin(xy)dx + x \sin(xy)dy = 0$

(c.) $-x^2 dy + y^2 dx = 0$

PP 36 A level surface $F(x, y, z) = 0$ has gradient vector field $\nabla F = \langle F_x, F_y, F_z \rangle$ which serves as a normal to the tangent plane of the given surface. Since $\omega_{\nabla F} = dF$ we need that $d\omega_{\nabla F} = d(dF) = 0$. Determine if the vector fields below are normal vector fields to some level surface in $\mathbb{R}^3$:

(a.) $\langle x, y, z \rangle$

(b.) $\langle y, 1, x \rangle$

(c.) $\langle xy^2 + z, x^2 y, -x \rangle$

PP 37 Recall the vector field $\vec{F} = \langle a, b, c \rangle$ corresponds to the **work form** $\omega_{\vec{F}} = adx + bdy + cdz$. Suppose $\vec{F} = \langle yz, x, 3, z^3 \rangle$ then write the formula for $\omega_{\vec{F}}$ and calculate $d\omega_{\vec{F}}$. Is $\vec{F}$ conservative on $\mathbb{R}^3$?

PP 38 Recall the vector field $\vec{F} = \langle a, b, c \rangle$ corresponds to the **flux form** $\Phi_{\vec{F}} = ady \wedge dz + bdz \wedge dx + cdx \wedge dy$. Suppose $\vec{F} = \langle x + y, y + z, z + x \rangle$ then write the formula for $\Phi_{\vec{F}}$ and calculate $d\Phi_{\vec{F}}$. If we were to calculate the flux of $\vec{F}$ through the unit-sphere then what would the result be ?

PP 39 (Orthogonal Trajectories) Find the orthogonal trajectories to the curve or family of curves described below:

(a.) $y = (x - c_1)^2$

(b.) $y = e^{c_1 x}$

(c.) $y = \frac{1+c_1 x}{1-c_1 x}$

(d.) $\sinh y = c_1 x$

(e.) $y^2 - x^2 = c_1 x^3$

PP 40 (Polar Curves) Consider a curve with polar equation $r = f(\theta)$. Let Ψ be the counterclockwise angle swept from the radial line to the tangent line along the curve. Show that $r\frac{d\theta}{dr} = \tan\Psi$. Then show that two polar curves are orthogonal if and only if $\tan\Psi_1 \tan\Psi_2 = -1$ at a point of intersection between curve C_1 and C_2. Use $r = f_1(\theta)$ for curve C_1 whereas $r = f_2(\theta)$ for C_2.

PP 41 (Orthogonal Trajectories to Polar Curves) Find the orthogonal trajectory for the curves described below (please use polar coordinates to formulate the answer)

(a.) $r = c_1(1 + \cos\theta)$

(b.) $r = \frac{c_1}{1+\cos\theta}$

(c.) $r = c_1 e^{\theta}$

PP 42 (Isogonal Families) A family of curves which intersects a given family of curves at an angle $\alpha \neq \pi/2$ are said to be **isogonal trajectories** of each other. If $\frac{dy}{dx} = f(x,y)$ describes a given family of curves then show its isogonal family are solutions of

$$\frac{dy}{dx} = \frac{f(x,y) \pm \tan\alpha}{1 \mp f(x,y)\tan\alpha}.$$

Then, find the isogonal family to $y = c_1 x$ at angle $\alpha = 30^o$.

PP 43 Find a Cartesian equation which is parametrized by the solution of the following system of differential equations:

$$\frac{dx}{dt} = -y \qquad \& \qquad \frac{dy}{dt} = 2x.$$

PP 44 An integral curve to a vector field $\vec{F} = \langle P, Q\rangle$ can be described parametrically as a path $t \mapsto \vec{\gamma}(t) = (x(t), y(t))$ for which $\vec{F}(\vec{\gamma}(t)) = \frac{d\vec{\gamma}}{dt}$. That is, $\langle P, Q\rangle = \langle dx/dt, dy/dt\rangle$. Parametrically we need to solve $\frac{dx}{dt} = P$ and $\frac{dy}{dt} = Q$. However, if we are only interested in describing the integral curve in Cartesian coordinates then we can eliminate t via the calculus

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{Q}{P}$$

thus finding an integral curve for a given vector field which depends only on x, y is as simple as solving the above first order ODEqn.

(a.) Find integral curves to the vector field $\vec{F}(x,y) = \langle x^2, y^3\rangle$.

(b.) Consider the vector field $\vec{F}(x,y) = \left\langle \frac{y}{(x-1)^2+y^2}, \frac{1-x}{(x-1)^2+y^2} \right\rangle$. Find the the level curve which serves as an integral curve for $\vec{F}$ through $P_o = (x_o, y_o) \neq (1,0)$.

(c.) Find the integral curves of the vector field $\vec{F} = \langle 1, e^{x^3} - 2y/x\rangle$. Please leave your answer explicitly in terms of y as a function of x.

PP 45 Suppose $F_f = -\alpha v^3$ is the net-force acting on a mass m in one-dimensional motion where the coordinate is denoted x (here α is an appropriate dimensional constant). Suppose $v = v_o$ and $x = x_o$ when $t = 0$ and calculate:

(a) velocity as function of time t

(b) velocity as a function of position x

PP 46 Let b be a positive constant. If a friction force of $F_f = -bv^4$ is applied to a mass m with initial position x_o and initial velocity v_o then find the velocity as a function of

(a.) time t,

(b.) position x.

PP 47 A net-force of $F = \alpha t - kv$ is placed on a mass m where k, α are constant. Find the velocity as an **explicit** function of time t. Assume that the particle undergoes one-dimensional motion with velocity v_o when $t = 0$.

PP 48 A large tank initially has 20 lbs of Koolaid mix added to 1000 *gallons* of water. Studies have shown that children will only drink the Koolaid when there is at least 0.005 *lbs* per gallon of water. You work for an incredibly lazy camp director who insists on adding the Koolaid to the 1000 *gallon* mixing tank only when the kids finally start rejecting the sadly weak Koolaid. Given that 25 *gallons* of pure water are added to the tank every day to make up for the 25 *gallons* of Koolaid drunk by the kids then when will you have to ask the camp director to add more mix ?

PP 49 A tank initially contains 100 gallons of water with 10lb of lemon drink mix. Then at $t = 0$ fresh water is added to the tank at 3 gallons per minute and at the same time 3 gallons are drained per minute from the tank. Assume the tank is well-mixed during this process. Find the lb's of lemon drink mix as a function of time. If you like your drink with a concentration of 1lb per 20 gallons then at what time should you drink from the drain?

PP 50 Suppose a mixing tank is well-stirred and contains y kilograms of salt at time t. Suppose pure water flows into the tank at a rate of 4 Liters per minute and salty water flows out at a rate of 2 Liters per minute. If the tank has 300 Liters of liquid and $300kg$ at time zero then write (but do not solve) the differential equation which describes the change in y.

PP 51 The radioactive lead isotope Pb-209 has a half-life of 3.3 hours. If 1 kilogram is initially present then how long will it take for there to be only 0.1 kilograms of radioactive lead remaining ?

PP 52 Show that the half-life of a radiactive substance is given by

$$t = \frac{(t_2 - t_1)\ln 2}{\ln(A_1/A_2)}$$

where $A_1 = A(t_1)$ and $A_2 = A(t_2)$ where $t_1 < t_2$.

PP 53 When a vertical beam of light passes through a transparent substance the rate at which its intensity I decreases is proportional to $I(y)$ where y represents the thickness of the medium in feet. In clear calm seawater the intensity 3 feet below the surface is 1/4 the initial intensity of the incident beam at the surface. What is the intensity of the beam 15 feet below the surface ?

PP 54 A thermometer is removed from a room where the air temperature is 70^o F to the outside where the temperature is 10^o F. After 0.5 minutes the thermometer reads 50^o F. What is the reading at $t = 1.0$ minutes ? How long will it take for the thermometer to reach 15^o F ?

PP 55 When a resistor R and inductor L are in series with a voltage source $\mathcal{E}$ then circuit analysis yields the differential equation:

$$L\frac{di}{dt} + Ri = \mathcal{E}$$

where i is the current flowing in the circuit. Given $\mathcal{E}(t) = V_o \sin \omega t$ and $i(0) = i_o$ find the current as a function of time t.

PP 56 Angular momentum of a body moving in some plane is given by $L = mr^2 \frac{d\theta}{dt}$ where r, θ serve as polar coordinates in the plane of motion. Assume the coordinates of the body are (r_1, θ_1) at $t = t_1$ and (r_2, θ_2) at $t = t_2$ where $t_1 < t_2$. **If L is constant then show that the area swept out by r is $A = L(t_2 - t_1)/2m$.** When the sun is taken to be at the origin and m represents a planet's mass then this proves Kepler's second law of planetary motion: the radius vector joining the sun sweeps out equal areas for equal intervals of time. **Bonus: prove L is constant in the context of the sun-planet system, you may assume $M_{sun} >> M_{planet}$. See Physics 231 for further relevant definitions.**

PP 57 Find the velocity of a mass m which is launched vertically with velocity v_o from a planet with mass M and radius R. Recall that the gravitational force is given by:

$$F = -\frac{GmM}{(R + y)^2}$$

if we assume the motion is directly vertical and y is the altitude of m. You may find the velocity as a function of y.

PP 58 Suppose a rocket car has an initial speed of v_o as it hurtles across a speedway in a remote desert. Suppose the driver opens a parachute which developes a retarding force proportional to the cube of the velocity; $F_f = -kv^3$. Find the velocity as:

(a.) a function of time,

(b.) a function of position x taking x_o as the initial position

PP 59 A ball is thrown verically and its motion is governed by the force of gravity $-mg$ and a friction force $F_f = -cv$ where $v = \dot{y}$ and c is a constant which is based on the size of the ball. Note: $F_f > 0$ when $v < 0$ and $F_f < 0$ when $v > 0$ hence the friction force acts opposite of the motion.

(a) find the velocity as a function of time and the initial velocity v_o

(b) find the position y as a function of time and the initial position y_o

(c) if $c = 0$ then what is the maximum height reached by the ball? Does it depend on the m?

(d) if $c \neq 0$ then what is the maximum height reached by the ball? Does it depend on the m?

PP 60 Suppose a $10kg$ block is pushed across a surface by a constant force of $10N$. As the block moves it gathers a gummy substance which results in a friction force of magnitude $F_o e^{kt}$ where $F_o = 1N$ and $k = 1/s$. Find the position x of the block as a function of time t. Assume that at time $t = 0$ the block is at $x = 0$. (feel free to use technology to do the numerical aspects here)

PP 61 A chain is coiled on the ground. One end is then lifted with constant force. Find the velocity.

PP 62 Suppose the RL-circuit has a voltage source which varies with time according to $\mathcal{E}(t) = V_o \cos(t)$. Find the current as a function of time and the initial current I_o. *hint: this is like an example in the notes, just replace the constant* $\mathcal{E}$ *with the sinuisoidal source* $\mathcal{E}(t) = V_o \cos(t)$

PP 63 Find a continuous function P such that $P(x) = a + \int_0^x t^2 P(t)\, dt$.

PP 64 Suppose f is a function such that $f(x+y) = f(x)f(y)$ for all $x, y \in \mathbb{R}$. Show that either f is the function which is identically zero on $\mathbb{R}$ or$f$ is an exponential function.

PP 65 Let $\star$ be the DEqn $y^2 \sin(x)dx + yf(x)dy = 0$. Find all functions f such that $\star$ is an exact DEqn.

PP 66 Explicitly solve $\dfrac{dy}{dx} = \dfrac{2x + 3x^2}{2y}$ given that the point $(1, -2)$ is on the solution.

PP 67 Solve $\dfrac{dy}{dx} - \dfrac{3}{x}y = x^3 e^x$

PP 68 Solve $\left(\frac{1}{x+1} + y\cos(xy)\right) dx + (e^{-y} + x\cos(xy))\, dy = 0$.

PP 69 Find orthogonal trajectories of $\frac{dy}{dx} = \frac{-x}{y}$.

PP 70 Find the velocity as a function of time t given that $v = v_o$ when $t = t_o$ and $F_{net} = -\beta v^2$ for a mass m.

PP 71 Solve $\frac{dP}{dt} = kP(P - C)$. Describe the possible solutions. You should find three disjoint types. Assume $k > 0$.

PP 72 Consider the following four differential equations:

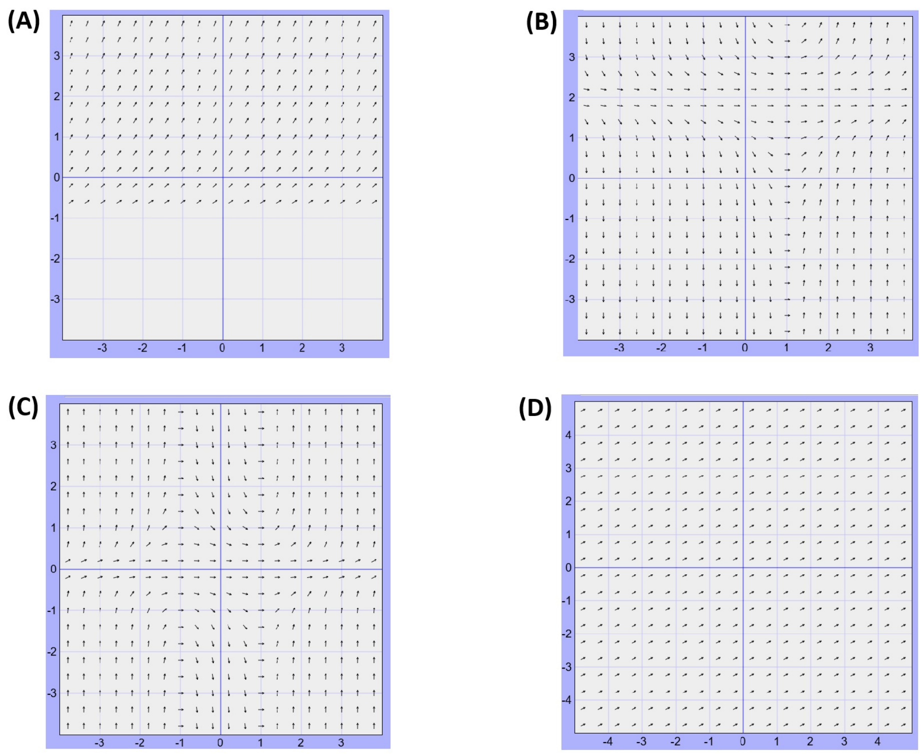
(I.) $\frac{dy}{dx} = 1/2$

(II.) $\frac{dy}{dx} = \sqrt{y+1}$

(III.) $\frac{dy}{dx} = (y-2)^2(x-1)$

(IV.) $\frac{dy}{dx} = (x^2-1) * y^2$

I used https://aeb019.hosted.uark.edu/pplane.html to generate the following direction fields. Match A,B,C,D with the corresponding I,II,III,IV.



Remark: the pplane tool is an easy way to explore the behavior of a given differential equation without going to the trouble of solving it. To visualize $\frac{dy}{dx} = f(x,y)$ I set $dx/dt = 1$ and $dy/dt = f(x,y)$. I didn't click on any points for the above plots, but if you do it traces out solutions. This can also illustrate solutions to systems of two autonomous ODEs, we'll get to that a bit later in the course.

Chapter 3

ordinary n-th order problem

To be honest, we would more properly title this chapter "the ordinary **linear** n-th order problem". Our focus is entirely in that direction here.

Our main approach in this chapter is to use **operator techniques**. In particular, we derive the operator calculus in Section 3.1. We show that the algebra of polynomials applies naturally to operators which are formed by polynomials in the derivative operator. Furthermore, we examine the *complexification technique* and see how it allows us to extract a pair of real solutions from a given complex solution. Section 3.1 is the computational machine which drives the rest of our work and at its heart is factoring. Many factoring techniques are illustrated throughout Section 3.1.

Next we turn to the general theory of linear ODEs. We define linear ODEs, describe the initial value problem and present the fundamental existence and uniqueness theorems in Section 3.2. Linear independence is justified with the help of the Wronskian and the insight of Abel's formula and we find linear combinations in the formulation of the general solution.

In Section 3.3 we solve the homogeneous constant coefficient problem. Then in Section 3.4 we see how to convert certain nonhomogeneous problems into homogeneous constant coefficient problems. This justifies the *method of undetermined coefficients*. Whenever the method of undetermined coefficients fails, we seek the deeper magic of *variation of parameters*. Section 3.5 gives proof that , given a fundamental solution set and an inhomogeneous term, variation of parameters can always solve the system up to a set of integrals which we may or may not be able to calculate in closed-form.

The remaining sections deal with two important techniques which have general application to problems which may not fall under the simple framework of earlier sections. In Section 3.6 we find that if we know one solution to an n-th order problem then we can use *reduction of order* to replace the given problem with an $(n-1)$-th order problem. Most often, I use this idea as it applies to the $n=2$ problem where we derive the excellent formula given in Equation 3.2. Then, Section 3.7 examines how factored operators naturally allow a repeated substitution solution. Unfortunately, it's usually harder to factor an operator than to solve the system so Section 3.7 is not always useful. Section 3.8 shows how the Cauchy Euler problem is equivalent to solving differential equations which arise from polynomials in the operator $T = xd/dx$. Finally, we study retarded springs and things in Section 3.9.1. We spend some effort to analyze when a force or source voltage maximally couples to a system which is retarded by friction or resistance. Once more, the beauty of the complexification technique shines bright.[1]

[1]Ideally I would like to add a subsection on the phasor technique, but my time is up for this term, my apologies, perhaps we will see the topic in-class.

3.1 operators and calculus

In this section we seek to establish the calculus of operators. All the results here are based on Calculus I, but the notation is probably new. First, let us discuss what we mean by an **operator**.

> If $\mathcal{F}$ is a set of functions then L is an operator on $\mathcal{F}$ if L is a mapping which sends y to $L[y]$ for each function $y \in \mathcal{F}$. In other words, an operator is a function-valued mapping of functions.

The details of $\mathcal{F}$ are varied, it could be the set of polynomials $\mathbb{R}[x]$, or continuous functions on $(C^0(\mathbb{R}))$, or continuously differentiable functions $(C^1(\mathbb{R}))$, or twice-continuously differentiable functions $(C^2(\mathbb{R}))$, or k-times continuously differentiable functions $(C^k(\mathbb{R}))$, or even infinitely-differentiable (smooth) functions $(C^\infty(\mathbb{R}))$. My examples thus far had real domain and codomain, they were functions on $\mathbb{R}$. Generally, we can also consider functions with real domain and complex codomain. We have need of both and I will explain the calculus of both in this section.

You know several operators already from calculus. For instance, the definite integral with variable bounds allow us to define an operator T by

$$T[f](x) = \int_0^x f(t)dt$$

here we map the function f to the new function $T[f]$ and we write $f \mapsto T[f]$ to express this process. Probably the most important is the **differentiation operator** $D = d/dx$.

$$f \mapsto D[f] = \frac{df}{dx}$$

Notice D takes in a function f and outputs a new function f'. In the same way, iterated differentiation gives an operator; D^k maps the function f to the function $f^{(k)}$:

$$f \mapsto D^k[f] = \frac{d^k f}{dx^k}$$

There are easier operators as well, let $L[f](x) = cf(x)$ for each x in the domain of f we call this **scalar multiplication**. It to defines an operator:

$$f \mapsto cf$$

If we have a function g then we can also define an operator $L[f](x) = g(x)f(x)$ for each x in the common domain of f and g. Once more, this defines an operator

$$f \mapsto gf$$

we usually call this **multiplication by** g. The examples above all share the interesting and important feature of **linearity**.

Definition 3.1.1. *linear operators*

> We say $T : \mathcal{F} \to \mathcal{F}$ is a **linear operator** if
>
> $$T(y_1 + y_2) = T(y_1) + T(y_2) \qquad \& \qquad T(cy) = cT(y)$$
>
> for all functions y_1, y_2, y and constants c. .

If you desire a non-linear operator those are easy enough to find:

Example 3.1.2. *Let $T[y] = y^2$ then $T[y_1 + y_2] = (y_1 + y_2)^2 = y_1^2 + 2y_1y_2 + y_2^2 = T[y_1] + T[y_2] + 2y_1y_2$. Thus T is not* **additive** *as $T[y_1 + y_2] \neq T[y_1] + T[y_2]$. Likewise, $T[cy] = (cy)^2 = c^2y^2 = c^2T[y]$ hence T is not* **homogeneous** *in that $T[cy] \neq cT[y]$.*

Now that we have a few common operators in mind it is good for us to define how we add, subtract and multiply operators. My apologies for this Machiavellian definition:

Definition 3.1.3. *operators, operator equality, new operators from old:*

Suppose $\mathcal{F}$ is a set of functions then $T : \mathcal{F} \to \mathcal{F}$ is an operator on $\mathcal{F}$. If T_1, T_2 are operators on $\mathcal{F}$ then $T_1 = T_2$ if and only if $T_1[f] = T_2[f]$ for all $f \in \mathcal{F}$. In addition, $T_1 + T_2$, $T_1 - T_2$ and T_1T_2 are defined by

$$(T_1 + T_2)[f] = T_1[f] + T_2[f] \quad \& \quad (T_1 - T_2)[f] = T_1[f] - T_2[f] \quad \& \quad (T_1T_2)[f] = T_1[T_2[f]]$$

If $g \in \mathcal{F}$ and T is an operator on $\mathcal{F}$ then gT and Tg are the operators defined by $(gT)[f] = gT[f]$ and $(Tg)[f] = T[f]g$ for all $f \in \mathcal{F}$. In addition, for $n \in \mathbb{N} \cup \{0\}$ we define T^n by $T^{n-1}T$ where $T^0 = 1$; that is $T^n[f] = T[T[\cdots[T[f]]\cdots]]$ where that is an n-fold composition. We are often interested in differentiation thus it is convenient to denote differentiation by D; $D[f] = f'$ for all $f \in \mathcal{F}$.

We usually assume $\mathcal{F}$ is a set of smooth functions on an interval. In fact, sorry to be sloppy, but we will not worry too much about what $\mathcal{F}$ is for most examples. Probably the most challenging aspect of the above definition is the multiplication of operators. Let us study an example to better understand the nuance of what is said above.

Example 3.1.4. *Let $T[f](x) = f'(x)$ and let $S[f](x) = xf'(x)$ for each $x \in \mathbb{R}$ and smooth function f. In other words, let $T = D$ and $S = xD$. Consider,*

$$(TS)[f](x) = T(S(f(x))) = D[xD[f]] = \frac{d}{dx}\left[x\frac{df}{dx}\right] = \frac{dx}{dx}\frac{df}{dx} + x\frac{d^2f}{dx^2} = D[f] + xD^2[f] = (D + xD^2)[f].$$

Since the calculation above holds for all f we find the **operator equation** *$TS = D + xD^2$. In other words, $DxD = D + xD^2$. On the other hand, if we compose the operators in the other order the calculation is not the same:*

$$(ST)[f](x) = S(T(f(x))) = xD[D[f]] = (xD^2)[f](x)$$

which reveals $ST = xD^2$ or we could write $xDD = xD^2$. Notice that $ST \neq TS$.

We can study the **commutator** of two operators S, T defined by $[S, T] = ST - TS$. Then the operators commute if and only if the commutator of the operators is zero. This is of great importance to quantum mechanics where we use operators to represent physically observable quantities like momentum and position. The Heisenberg uncertainty principle famously says it is impossible to measure both position and momentum with perfect precision. It turns out this is downstream of the fact that the operators of position and momentum do not commute. A theorem of quantum mechanics states that for two observables to be simultaneously observed their corresponding operators must have zero commutator.

Example 3.1.5. *Let $T = D + 3$ and $S = D - 2$ where $D = d/dx$. Let f be a function of x,*

$$(TS)[f](x) = T(S(f(x))) = (D + 3)\left[\frac{df}{dx} - 2f\right] = \frac{d}{dx}\left[\frac{df}{dx} - 2f\right] + 3\left[\frac{df}{dx} - 2f\right]$$

which is easier to write using the D-notation, and we may omit x without danger of ambiguity in this context,

$$(TS)[f] = D[D[f] - 2f] + 3(D[f] - 2f) = D^2[f] - 2D[f] + 3D[f] - 6f = (D^2 + D - 6)[f].$$

Therefore, we find the operator equation $TS = (D+3)(D-2) = D^2 + D - 6$. The reader will not be surprised to learn an entirely similar calculation can be made to demonstrate that $ST = (D-2)(D+3) = D^2 + D - 6$. Hence $ST = TS$. These operators **do** *commute.*

The example above illustrates that polynomials in the differentiation operator are especially nice. We should pause to give a general definition of what we mean by the *polynomial of an operator*. This will be important for future work in this course.

Definition 3.1.6. *polynomial of an operator*

Suppose $\mathcal{F}$ is a set of functions and $T : \mathcal{F} \to \mathcal{F}$ is an operator on $\mathcal{F}$. Let $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial then we define

$$P(T) = a_n T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0.$$

In Linear Algebra we prove that the sum, difference and composition of linear operators is once more a linear operator. It follows that if T is a linear operator than so is $P(T)$. These claims are easy to verify in particular examples.

Example 3.1.7. *Let $T[y] = y'' + 3y' + 2y$ that is $T = D^2 + 3D + 2$. Observe,*

$$T[y_1 + y_2] = (y_1 + y_2)'' + 3(y_1 + y_2)' + 2(y_1 + y_2) = y_1'' + y_2'' + 3y_1' + 3y_2' + 2y_1 + 2y_2 = T[y_1] + T[y_2].$$

Likewise,

$$T[cy] = (cy)'' + 3(cy)' + 2(cy) = c(y'' + 3y' + 2y) = cT[y].$$

So we have shown T is a linear operator. Consider also,

$$(D + 1)(D + 2)[y] = (D + 1)[y' + 2y] = (y' + 2y)' + y' + 2y = y'' + 3y' + 2y = (D^2 + 3D + 2)[y]$$

Thus $(D + 1)(D + 2) = D^2 + 3D + 2$.

This is not an accident.

Theorem 3.1.8. *polynomials in the derivative operator factor*

Suppose $p(x) = f(x)g(x)$ where $f(x), g(x)$ are polynomials. Then $p(D) = f(D)g(D)$ and $f(D)g(D) = g(D)f(D)$.

Proof: Suppose $p(x) = f(x)g(x)$ where $g(x) = a_0 + a_1 x + \cdots + a_k x^k$ and $f(x) = b_0 + b_1 x + \cdots + b_m x^m$.

$$\begin{aligned}
(f(D)g(D))[y] &= f(D)(g(D)[y]) \\
&= (b_0 + b_1 D + \cdots + b_m D^m)(g(D)[y]) \\
&= b_0 g(D)[y] + b_1 D(g(D)[y]) + \cdots + b_m D^m(g(D)[y]) \\
&= b_0 \left(a_0 y + a_1 D[y] + \cdots + a_k D^k[y]\right) + b_1 D \left(a_0 y + a_1 D[y] + \cdots + a_k D^k[y]\right) + \cdots \\
&\quad \cdots + b_m D^m \left(a_0 y + a_1 D[y] + \cdots + a_k D^k[y]\right) \\
&= \big(b_0 a_0 + b_0 a_1 D + \cdots + b_0 a_k D^k + b_1 a_0 D + b_1 a_1 D^2 + \cdots + b_1 a_k D^{k+1} + \cdots \\
&\quad \cdots + b_m a_0 D^m + b_m a_1 D^{m+1} + \cdots + b_m a_k D^{m+k}\big)[y] \\
&= p(D)[y]
\end{aligned}$$

since $p(x) = f(x)g(x) = (b_0 + b_1 x + \cdots + b_m x^m)(a_0 + a_1 x + \cdots + a_k x^k)$ which multiplies by the usual polynomial algebra to give that:

$$\begin{aligned}
p(x) = f(x)g(x) &= b_0 a_0 + b_0 a_1 x + \cdots + b_0 a_k x^k + b_1 a_0 x + b_1 a_1 x^2 + \cdots + b_1 a_k x^{k+1} + \cdots \\
&\quad \cdots + b_m a_0 x^m + b_m a_1 x^{m+1} + \cdots + b_m a_k x^{m+k}.
\end{aligned}$$

Finally, if $p(x) = f(x)g(x)$ then $p(x) = g(x)f(x)$ as polynomial multiplication commutes and we likewise find $p(D) = f(D)g(D) = g(D)f(D)$. $\square$

The calculation above equally well applies to a polynomial in a linear transformation T since T commutes with itself just like D. For instance, the Cauchy Euler problem studied in Section 3.8 amounts to working with polynomials in the operator $T = x\frac{d}{dx}$. It follows we can solve a Cauchy Euler problem in much the same way as we solve a constant coefficient problem. The constant coefficient problem is based on a polynomial in $D = d/dx$ whereas the Cauchy Euler problem is based on the equidimensional operator $T = xd/dx$.

3.1.1 on the derivation of real solutions for homogenous ODEs

In this subsection we study how to find real solutions of $(D-\lambda)^k[y]=0$ for $k=1,2,\dots$.

(1.) $(k=1)$ If $(D-\lambda)[y]=0$ then we require $\frac{dy}{dx}-\lambda y=0$ hence

$$\frac{dy}{dx}=\lambda y \quad\Rightarrow\quad \frac{dy}{y}=\lambda dx \quad\Rightarrow\quad \ln|y|=\lambda x+c \quad\Rightarrow\quad y_1=c_1e^{\lambda x}$$

serves as a solution. For future reference in this section, we set $c_1=1$ for convenience; $\boxed{y_1=e^{\lambda x}}$.

(2.) $(k=2)$ If $(D-\lambda)^2[y]=0$. Let $z=(D-\lambda)[y]$ then we face $(D-\lambda)[z]=0$ hence $z=e^{\lambda x}$ from our work in $k=1$. Thus, using the integrating factor technique we find that:

$$(D-\lambda)[y]=e^{\lambda x} \quad\Rightarrow\quad \frac{dy}{dx}-\lambda y=e^{\lambda x} \quad\Rightarrow\quad e^{-\lambda x}\frac{dy}{dx}-\lambda e^{-\lambda x}y=e^{-\lambda x}e^{\lambda x}=1 \quad\Rightarrow\quad \frac{d}{dx}\left[e^{-\lambda x}y\right]=1.$$

Consequently, integrating yields $e^{-\lambda x}y=x+c$ and so $y=xe^{\lambda x}+ce^{\lambda x}$. Let us define $\boxed{y_2=xe^{\lambda x}}$ and notice that both $y_1=e^{\lambda x}$ and $y_2=xe^{\lambda x}$ serve as solutions of $(D-\lambda)^2[y]=0$.

(3.) $(k=3)$ If $(D-\lambda)^3[y]=0$. Let $z=(D-\lambda)[y]$ then we face $(D-\lambda)^2[z]=0$ hence $z=xe^{\lambda x}$ from our work in $k=2$. Thus, using the integrating factor technique we find that:

$$(D-\lambda)[y]=xe^{\lambda x} \quad\Rightarrow\quad \frac{dy}{dx}-\lambda y=xe^{\lambda x} \quad\Rightarrow\quad e^{-\lambda x}\frac{dy}{dx}-\lambda e^{-\lambda x}y=xe^{-\lambda x}e^{\lambda x}=x \quad\Rightarrow\quad \frac{d}{dx}\left[e^{-\lambda x}y\right]=x.$$

Integrating we find $e^{-\lambda x}y=x^2/2$ and so $y=\frac{1}{2}x^2e^{\lambda x}$. However, $0=(D-\lambda)^2\left[\frac{1}{2}x^2e^{\lambda x}\right]=\frac{1}{2}(D-\lambda)^2\left[x^2e^{\lambda x}\right]$ thus we set $\boxed{y_3=x^2e^{\lambda x}}$ since we don't need the fraction to solve the problem.

(4.) $(k=4)$. Very similar calculations reveal $y_4=x^3e^{\lambda x}$ solve $(D-\lambda)^4[y]=0$.

I hope the calculations above serve to convince the reader that the following theorem is plausible. For the math majors, I'll offer an explicit proof below.

Theorem 3.1.9. *repeated real root solution set*

> Let $\lambda\in\mathbb{R}$. Then $e^{\lambda x}, xe^{\lambda x}, x^2e^{\lambda x},\dots,x^{n-1}e^{\lambda x}$ are solutions of $(D-\lambda)^n[y]=0$.

Proof: Let $\lambda\in\mathbb{R}$. We claim $e^{\lambda x}, xe^{\lambda x}, x^2e^{\lambda x},\dots,x^{n-1}e^{\lambda x}$ are solutions of $(D-\lambda)^n[y]=0$ for every $n\in\mathbb{N}$. Observe $n=1$ holds true since

$$(D-\lambda)^n[e^{\lambda x}]=D[e^{\lambda x}]-\lambda e^{\lambda x}=\lambda e^{\lambda x}-\lambda e^{\lambda x}=0.$$

Inductively suppose that $e^{\lambda x}, xe^{\lambda x}, x^2e^{\lambda x},\dots,x^{n-1}e^{\lambda x}$ are solutions of $(D-\lambda)^n[y]=0$. Since $(D-\lambda)^n[y]=0$ for $y=e^{\lambda x}, xe^{\lambda x}, x^2e^{\lambda x},\dots,x^{n-1}e^{\lambda x}$ we find

$$(D-\lambda)^{n+1}[y]=(D-\lambda)((D-\lambda)^n[y])=(D-\lambda)(0)=0.$$

It remains to show $x^ne^{\lambda x}$ solves $(D-\lambda)^{n+1}[y]=0$. Consider,

$$\begin{aligned}(D-\lambda)^{n+1}[x^ne^{\lambda x}]&=(D-\lambda)^n(D-\lambda)[x^ne^{\lambda x}]\\&=(D-\lambda)^n(nx^{n-1}e^{\lambda x}+\lambda x^ne^{\lambda x}-\lambda x^ne^{\lambda x}]\\&=n(D-\lambda)^n[x^{n-1}e^{\lambda x}]\\&=0\end{aligned}$$

by the induction hypothesis. Thus $y=e^{\lambda x}, xe^{\lambda x}, x^2e^{\lambda x},\dots,x^ne^{\lambda x}$ solve $(D-\lambda)^{n+1}[y]=0$ and we conclude the claim holds true for all $n\in\mathbb{N}$ by proof by mathematical induction. $\square$

Example 3.1.10. $(D^2-6D+9D)[y]=(D-3)^2[y]=0$ *has solutions* $y_1=e^{3x}$ *and* $y_2=xe^{3x}$.

Example 3.1.11. $(D^3+3D^2+3D+1)[y]=(D+1)^3[y]=0$ *has solutions* $y=e^{-x},xe^{-x},x^2e^{-x}$.

Example 3.1.12. $(D^4-4D^3+6D^2-4D+1)[y]=(D-1)^4[y]=0$ *has solutions* $y=e^x,xe^x,x^2e^x,x^3e^x$.

I am using Pascal's triangle and the binomial theorem to factor the rather special polynomials in the above examples. We need a bit more theory to extend to more interesting examples.

Theorem 3.1.13.

Suppose $L_1,L_2,\dots,L_k$ are linear transformations. If $L_k[y]=0$ then $(L_1L_2\cdots L_k)[y]=0$.

Proof: notice linearity of L gives $L[0]=L[0+0]=L[0]+L[0]$ hence $L[0]=0$. Furthermore, the composition of linear transformations is once more a linear transformation hence $L=L_1L_2\cdots L_{k-1}$ is a linear transformation. Suppose $L_k[y]=0$ and observe

$$(L_1L_2\cdots L_k)[y]=L(L_k[y])=L(0)=0. \qquad \square$$

Let me illustrate how the theorem above expands our analysis.

Example 3.1.14. $(D^2-3D+2)[y]=(D-1)(D-2)[y]=0$ *has solution* $y_1=e^{2x}$. *However, we could just as well write* $(D^2-3D+2)[y]=(D-2)(D-1)[y]=0$ *and hence find solution* $y_2=e^x$.

Corollary 3.1.15.

If $L_1,L_2,\dots,L_k$ are commuting linear operators for which $L_j[y]=0$ for some $1\le j\le k$ then $(L_1L_2\cdots L_k)[y]=0$.

Proof: Suppose the operators $L_1,\dots,L_k$ commute and $L_j[y]=0$ for some $1\le j\le k$. Observe $L_1L_2\cdots L_k=L_1L_2\cdots L_{j-1}L_{j+1}\cdots L_kL_j$ and apply Theorem 3.1.13. $\square$
I hope you can forgive me for stating the theorem above in greater generality than our typical application. For the most part we will focus on operators formed by polynomials in the derivative operator $D=d/dx$ or $D=d/dt$. Let us appreciate the depth of the above result.

Example 3.1.16. *Here is a nice factoring technique when it's possible:*

$$D^4-3D^2+2=(D^2-1)(D^2-2)=(D+1)(D-1)(D+\sqrt{2})(D-\sqrt{2})$$

Consequently,

$$(D^4-3D^2+2)[y]=(D+1)(D-1)(D+\sqrt{2})(D-\sqrt{2})[y]=0$$

has solutions $y_1=e^{-x},y_2=e^x,y_3=e^{-x\sqrt{2}},y_4=e^{x\sqrt{2}}$.

Do you know how to complete the square for an expression ? It's time to learn if you don't.

Example 3.1.17. $D^2+4D+1=(D+2)^2-4+1=(D+2)^2-3=(D+2+\sqrt{3})(D+2-\sqrt{3})$. *Consequently,*

$$(D^2+4D+1)[y]=(D+2+\sqrt{3})(D+2-\sqrt{3})[y]=0$$

thus we find solutions $y_1=e^{(-2-\sqrt{3})x}$ *and* $y_2=e^{(-2+\sqrt{3})x}$

Example 3.1.18. *Let me illustrate completing the square once more, but now with fractions:*

$$D^2-D-13=\left(D-\frac{1}{2}\right)^2-\frac{1}{4}-13=\left(D-\frac{1}{2}\right)^2-\frac{53}{4}=\left(D-\frac{1}{2}+\sqrt{\frac{53}{4}}\right)\left(D-\frac{1}{2}-\sqrt{\frac{53}{4}}\right)$$

Consequently,

$$(D^2-D-13)[y]=\left(D-\frac{1-\sqrt{53}}{2}\right)\left(D-\frac{1+\sqrt{53}}{2}\right)[y]=0$$

thus we find solutions $y_1=exp\left(\left[\frac{1-\sqrt{53}}{2}\right]x\right)$ *and* $y_2=exp\left(\left[\frac{1+\sqrt{53}}{2}\right]x\right)$

Example 3.1.19. *Consider* $(D^4 - 10D^2 + 24)[y] = 0$. *Observe,*

$$\begin{aligned} D^4 - 10D^2 + 24 &= (D^2 - 5)^2 - 1 \\ &= (D^2 - 5 + 1)(D^2 - 5 - 1) \\ &= (D^2 - 4)(D^2 - 6) \\ &= (D + 2)(D - 2)(D + \sqrt{6})(D - \sqrt{6}). \end{aligned}$$

Thus $(D^4 - 10D^2 + 24)[y] = 0$ *has solutions* $y_1 = e^{-2x}, y_2 = e^{2x}, y_3 = e^{-x\sqrt{6}}, y_4 = e^{x\sqrt{6}}$.

I have been careful to avoid polynomials which have irreducible quadratic factors in this section. Generally we have no such freedom. We must face the difficulty. Hence we study complexification of operators next.

3.1.2 complex-valued functions of a real variable

Given a linear operator T on a set of real-valued functions of a real variable we form the **complexification** of $T_{\mathbb{C}}$ which acts on complex-valued functions of a real variable by the following rule:

$$T_{\mathbb{C}}[u + iv] = T[u] + iT[v].$$

In practice, we just write T for the complexification and hence $T[u + iv] = T[u] + iT[v]$. A good example of this construction is given by the derivative introduced below. It is the complexification of the standard real derivative:

Definition 3.1.20.

Suppose f is a function from an interval $I \subseteq \mathbb{R}$ to the complex numbers $\mathbb{C}$. In particular, suppose $f(t) = u(t) + iv(t)$ where $i^2 = -1$ we say $Re(f) = u$ and $Im(f) = v$. Furthermore, define

$$\frac{df}{dt} = \frac{du}{dt} + i\frac{dv}{dt} \qquad \& \qquad \int f(t)\,dt = \int u\,dt + i\int v\,dt.$$

Higher derivatives are similarly defined.

Example 3.1.21. *Let* $f(t) = \cos(t) + ie^t$. *In this case* $Re(f) = \cos(t)$ *and* $Im(f) = e^t$. *Note,*

$$\frac{df}{dt} = \frac{d}{dt}\Big(\cos(t) + ie^t\Big) = -\sin(t) + ie^t.$$

I invite the reader to verify the following properties for complex-valued functions f, g:

$$\frac{d}{dt}(f+g) = \frac{df}{dt} + \frac{dg}{dt} \qquad \& \qquad \frac{d}{dt}(cf) = c\frac{df}{dt} \qquad \& \qquad \frac{d}{dt}(fg) = \frac{df}{dt}g + f\frac{dg}{dt}$$

How to prove the product rule ? It's just a calculation. If $f = f_1 + if_2$ and $g = g_1 + ig_2$ then

$$fg = (f_1 + if_2)(g_1 + ig_2) = f_1g_1 - f_2g_2 + i(f_1g_2 + f_2g_1)$$

Let $df/dt = f'$ etc. There are four product rules, for the real component function products $f_1g_1, f_2g_2, f_1g_2, f_2g_1$

$$\begin{aligned} (fg)' &= f_1'g_1 + f_1g_1' - f_2'g_2 - f_2g_2' + i(f_1'g_2 + f_1g_2' + f_2'g_1 + f_2g_1') \\ &= (f_1' + if_2')(g_1 + ig_2) + (f_1 + if_2)(g_1' + ig_2') \\ &= f'g + fg'. \end{aligned}$$

The calculus of complex-valued functions of a real variable follows as a consequence of the corresponding calculus for functions on $\mathbb{R}$ paired with the algebraically simple properties of complex multiplication and addition. Note that the constant c can be complex in the property above and the multiplications between

f and g are multiplications of complex values. It turns out we can extend all the polynomial algebra of the derivative operator to the complexification of the derivatives. In short, all the arguments which we used in the real case equally well apply here since the complexified derivative D has the properties

$$D[f+g] = D[f] + D[g] \qquad \& \qquad D[cf] = cD[f] \qquad \& \qquad D[fg] = D[f]g + fD[g]$$

for complex-valued functions f, g. Each complex derivative contains two real derivatives. This means if we solve a complex differential equation then we may be able to find a pair of real solutions. This is certainly true in the context of the complexification of a real polynomial operator in $D = d/dx$. Let us record this as a theorem:

Theorem 3.1.22.

Let $f(x), g(x) \in \mathbb{R}[x]$ define $p(x) = f(x)g(x)$ and suppose T is a real linear operator then the complexification of T has $P(T) = f(T)g(T)$ and

$$P(T)[y+iz] = P(T)[y] + iP(T)[z]$$

for real functions y, z. In particular, $P(T)[y+iz] = 0 \;\Leftrightarrow\; P(T)[y] = 0 \;\&\; P(T)[z] = 0$.

Now, in order to make good use of the theorem above we need to study the complex generalization of the exponential function.

Definition 3.1.23. *complex exponential function.*

We define $exp : \mathbb{C} \to \mathbb{C}$ by the following formula: $e^{x+iy} = e^x\big(\cos y + i\sin y\big)$ where $x, y \in \mathbb{R}$.

We can show $exp(z+w) = exp(z)exp(w)$. Suppose that $z = x+iy$ and $w = a+ib$ where $x, y, a, b \in \mathbb{R}$,

$$\begin{aligned}
exp(z+w) &= exp(x+iy+a+ib) \\
&= exp(x+a+i(y+b)) \\
&= e^{x+a}\big(\cos(y+b) + i\sin(y+b)\big) && \text{defn. of complex exp.} \\
&= e^{x+a}\big(\cos y\cos b - \sin y \sin b + i[\sin y\cos b + \sin b\cos y]\big) && \text{adding angles formulas} \\
&= e^{x+a}\big(\cos y + i\sin y\big)\big(\cos b + i\sin b\big) && \text{algebra} \\
&= e^x e^a\big(\cos y + i\sin y\big)\big(\cos b + i\sin b\big) && \text{law of exponents} \\
&= e^{x+iy}e^{a+ib} && \text{defn. of complex exp.} \\
&= exp(z)exp(w).
\end{aligned}$$

In Math 331 we spend considerable effort to deal with the fact that $e^{z+2\pi ik} = e^z$ for any $k \in \mathbb{Z}$. We can only define an inverse locally, this is why the complex logarithm is a fascinating object, far more nuanced that the real natural log.

Proposition 3.1.24. *Let $\lambda = \alpha + i\beta$ for real constants α, β. We have:*

$$\frac{d}{dt}\big[e^{\lambda t}\big] = \lambda e^{\lambda t}.$$

Proof: direct calculation.

$$\begin{aligned}
\frac{d}{dt}\big[e^{\alpha t + i\beta t}\big] &= \frac{d}{dt}\big[e^{\alpha t}(\cos(\beta t) + i\sin(\beta t))\big] \\
&= \alpha e^{\alpha t}\cos(\beta t) - \beta e^{\alpha t}\sin(\beta t) + i\alpha e^{\alpha t}\sin(\beta t) + i\beta e^{\alpha t}\cos(\beta t) \\
&= \alpha e^{\alpha t}(\cos(\beta t) + i\sin(\beta t)) + i\beta e^{\alpha t}(\cos(\beta t) + i\sin(\beta t)) \\
&= (\alpha + i\beta)e^{\alpha t}(\cos(\beta t) + i\sin(\beta t)) \\
&= \lambda e^{\lambda t}. \qquad \square
\end{aligned}$$

This is a beautiful result. Let's examine how it works in an example.

Example 3.1.25. *Let $f(t) = e^{(2+i)t}$. In this case $f(t) = e^{2t}(\cos(t) + i\sin(t))$ thus $Re(f) = e^{2t}\cos(t)$ and $Im(f) = e^{2t}\sin(t)$. Note,*

$$\frac{df}{dt} = \frac{d}{dt}\left(e^{(2+i)t}\right) = (2+i)e^{(2+i)t}.$$

Expanding $(2+i)e^{2t}(\cos(t)+i\sin(t)) = 2e^t\cos(t) - e^t\sin(t) + i(2e^{2t}\sin(t) + e^{2t}\cos(t))$. Which is what we would naturally obtain via direct differentiation of $f(t) = e^{2t}\cos(t) + ie^{2t}\sin(t)$. Obviously the complex notation hides many details.

Example 3.1.26. *Note $D^2 + 1 = D^2 - i^2 = (D+i)(D-i)$ thus $(D^2+1)[y] = 0$ has complex solutions $z_1 = e^{ix}$ and $z_2 = e^{-ix}$. Notice $e^{ix} = \cos x + i\sin x$ thus*

$$0 = (D^2+1)[\cos x + i\sin x] = (D^2+1)[\cos x] + i(D^2+1)[\sin x]$$

which indicates $(D^2+1)[y] = 0$ has real solutions $y_1 = \cos x$ and $y_2 = \sin x$. Note, the conjugate solution $e^{-ix} = \cos x - i\sin x$ yields the real solutions $y_1 = \cos x$ and $-y_2 = -\sin x$.

Evidentally, we need only study one complex root of each conjugate pair since the functions associated with each root are the same up to a sign.

Example 3.1.27. *Note $D^2 - 4D + 13 = (D-2)^2 + 9 = (D-2-3i)(D-2+3i)$ thus $(D^2-4D+13)[y] = 0$ has complex solutions $z_1 = e^{(2+3i)x}$ and $z_2 = e^{(2-3i)x}$. Notice the complex solution*

$$e^{(2+3i)x} = e^{2x}\left(\cos(3x) + i\sin(3x)\right)$$

yields the real solutions $y_1 = e^{2x}\cos(3x)$ and $y_2 = e^{2x}\sin(3x)$ for $(D^2 - 4D + 13)[y] = 0$.

Complex roots of real polynomials always come in conjugate pairs. We might as well treat such pairs jointly. It will save us the trouble of factoring complex polynomials in many examples. Instead, we can just read off the solution via the following theorem:

Theorem 3.1.28. *complex conjugate pair solution set*

Let $\lambda = \alpha + i\beta \in \mathbb{C}$ where $\beta \neq 0$. Then for $n \in \mathbb{N}$,

$$e^{\alpha x}\cos(\beta x), e^{\alpha x}\sin(\beta x), xe^{\alpha x}\cos(\beta x), xe^{\alpha x}\sin(\beta x), \dots, x^{n-1}e^{\alpha x}\cos(\beta x), x^{n-1}e^{\alpha x}\sin(\beta x)$$

are real solutions of $\left((D-\alpha)^2 + \beta^2\right)^n[y] = 0$.

Proof: observe that $(D-\alpha)^2 + \beta^2 = (D - \alpha - i\beta)(D - \alpha + i\beta)$. Thus, using $\lambda = \alpha + i\beta$ and $\lambda^* = \alpha - i\beta$, $(D-\alpha)^2 + \beta^2 = (D-\lambda)(D-\lambda^*)$. Hence,

$$\left((D-\alpha)^2 + \beta^2\right)^n = (D-\lambda)^n(D-\lambda^*)^n$$

and we find $\left((D-\alpha)^2 + \beta^2\right)^n[y] = 0$ has complex solutions

$$z_1 = e^{\lambda x}, z_2 = xe^{\lambda x}, \dots, z_n = x^{n-1}e^{\lambda x}.$$

However, $e^{\lambda x} = e^{\alpha x}e^{+i\beta x} = e^{\alpha x}\cos(\beta x) + ie^{\alpha x}\sin(\beta x)$ thus

$$z_1 = e^{\alpha x}\cos(\beta x) + ie^{\alpha x}\sin(\beta x), z_2 = xe^{\alpha x}\cos(\beta x) + ixe^{\alpha x}\sin(\beta x), \dots, z_n = x^{n-1}e^{\alpha x}\cos(\beta x) + ix^{n-1}e^{\alpha x}\sin(\beta x).$$

But, we know the real and imaginary parts of complex solutions of $\left((D-\alpha)^2 + \beta^2\right)^n[y] = 0$ are separately real solutions of $\left((D-\alpha)^2 + \beta^2\right)^n[y] = 0$ hence the theorem follows. $\square$

Example 3.1.29. *$D^2 + 4D + 5 = (D+2)^2 + 1$ thus identify $\alpha = -2$ and $\beta = 1$ and $(D^2 + 4D + 5)[y] = 0$ has solutions $y_1 = e^{-2x}\cos x$ and $y_2 = e^{-2x}\sin x$.*

Example 3.1.30. *$(D^2+1)^2[y]=0$ has solutions $y_1=\cos x, y_2=\sin x, y_3=x\cos x, y_4=x\sin x$ stemming from $\lambda=i$ with $n=2$ where we've identified $\alpha=0$ and $\beta=1$.*

Finally, to conclude, we may also face problems where there is a mixture of real and complex roots. Let's examine a pair of examples:

Example 3.1.31. *$D^4-5D^2-36=(D^2-9)(D^2+4)=(D+3)(D-3)(D^2+4)$ thus $(D^4-5D^2-36)[y]=0$ has real solutions $e^{-3x}, e^{3x}, \cos(2x), \sin(2x)$.*

Example 3.1.32. *$D^4-20151121=(D^2-4489)(D^2+4489)=(D-67)(D+67)(D^2+67^2)$ thus $(D^4-20151121)[y]=0$ has real solutions $e^{67x}, e^{-67x}, \cos(67x), \sin(67x)$.*

We've simply combined our methods up to this point. Observe we're always able to produce n-distinct solutions of $p(D)[y]=0$ for an n-th order polynomial $p(D)$. The next section will show us how to take these n-solutions and use them to formulate general solutions for linear ODEs of the form $p(D)[y]=0$.

3.2 linear differential equations

We say I is an **interval** iff $I=(a,b),[a,b),(a,b],[a,b],[a,\infty),(a,\infty),(-\infty,a],(-\infty,a),(-\infty,\infty)$.

Definition 3.2.1. *n-th order linear differential equation*

Let I be an interval of real numbers. Let $a_0, a_1, \dots, a_n, f$ be real-valued functions on I such that $a_n(x)\neq 0$ for all $x\in I$. We say

$$a_n\frac{d^ny}{dx^n}+a_{n-1}\frac{d^{n-1}y}{dx^{n-1}}+\cdots+a_1\frac{dy}{dx}+a_0y=f \tag{3.1}$$

is an **n-th order linear differential equation** on I with **coefficient functions** $a_0, a_1, \dots, a_n$ and **forcing** function f. If $f(x)=0$ for all $x\in I$ then we say the differential equation is **homogeneous**. However, if $f(x)\neq 0$ for at least one $x\in I$ then the differential equation is said to be **nonhomogeneous**.

In the prime notation Equation 3.1 is denoted:

$$a_ny^{(n)}+a_{n-1}y^{(n-1)}+\cdots+a_1y'+a_0y=f$$

If the independent variable was denoted by t then we could emphasize that by writing

$$a_ny^{(n)}(t)+a_{n-1}y^{(n-1)}(t)+\cdots+a_1y'(t)+a_0y(t)=f(t).$$

Typically we either use x or t as the independent variable in this course. We denote differentiation as an **operator** D and, depending on the context, either $D=d/dx$ or $D=d/dt$. In this operator notation we can write Equation 3.1 as

$$a_nD^n[y]+a_{n-1}D^{n-1}[y]+\cdots+a_1D[y]+a_0y=f$$

or, introducing

$$L=a_nD^n+a_{n-1}D^{n-1}+\cdots+a_1D+a_0$$

we can express Equation 3.1 concisely as $L[y]=f$. We can show L is a **linear operator**:

Definition 3.2.2. *linear operator*

We say L is a linear operator if for any pair of functions y_1, y_2 and constant c we have

$$L[y_1+y_2]=L[y_1]+L[y_2] \qquad \& \qquad L[cy_1]=cL[y_1].$$

A function is said to be **smooth** if you can take arbitarily many derivatives of the function. In fact, if all the coefficient functions $a_n, \dots, a_1, a_0$ are smooth then $L=a_nD^n+a_{n-1}D^{n-1}+\cdots+a_1D+a_0$ defines a **smooth differential operator** in the sense that L maps a smooth function y to a smooth function $L[y]$.

Theorem 3.2.3. *linear combination of homogeneous solutions is a homogeneous solution.*

Let L be a linear operator and suppose y_1 and y_2 solve $L[y] = 0$. If c_1, c_2 are constants then $y = c_1 y_1 + c_2 y_2$ is a solution of $L[y] = 0$.

Proof: suppose $L[y_1] = 0$ and $L[y_2] = 0$ where L is a linear operator. Let c_1, c_2 be constants and calculate:

$$L[c_1 y_1 + c_2 y_2] = c_1 L[y_1] + c_2 L[y_2] = c_1(0) + c_2(0) = 0.$$

Thus $y = c_1 y_1 + c_2 y_2$ is a solution of $L[y] = 0$. $\square$

The result above extends to the case of n-homogenous solutions. If $L[y_j] = 0$ for $j = 1, 2, \dots, n$ then

$$L[c_1 y_1 + c_2 y_2 + \cdots + c_n y_n] = 0.$$

When solving a homogeneous differential equation $L[y] = 0$ we can always combine solutions by forming a linear combination of solutions.

In contrast, if we have y_1, y_2 solutions of a non-homogeneous differential equation $L[y] = f$ then $y = y_1 + y_2$ is not a solution since

$$L[y_1 + y_2] = L[y_1] + L[y_2] = f + f = 2f \neq f$$

for $f \neq 0$. Likewise, for $c \neq 1$, if $L[y] = f$ then $L[cy] = cL[y] = cf \neq f$ for $f \neq 0$. We can say something positive about non-homogeneous problems. In fact, the following theorem is at the heart of why linear differential equations are simple to analyze:

Theorem 3.2.4. *superposition principle.*

Let L be a linear operator and suppose y_1 has $L[y_1] = f_1$ whereas $L[y_2] = f_2$. If a, b are constants then $y = ay_1 + by_2$ is a solution of $L[y] = af_1 + bf_2$.

Proof: suppose $L[y_1] = f_1$ and $L[y_2] = f_2$ where L is a linear operator. Let a, b be constants and calculate:

$$L[ay_1 + by_2] = aL[y_1] + bL[y_2] = af_1 + bf_2.$$

Thus $y = ay_1 + by_2$ is a solution of $L[y] = af_1 + bf_2$. $\square$

Notice this means we can solve $L[y] = f_1 + f_2$ by solving $L[y] = f_1$ and $L[y] = f_2$ separately. Physically speaking, we can think of $L[y] = f$ as a physical system subject to force f. The principle of superposition indicates that if the net-force is $f_1 + f_2$ then the motion $y_1 + y_2$ can be understood the sum of motion y_1 due to force f_1 and the motion y_2 due to f_2. We make this precise in our study of the retarded spring problem towards the end of this Chapter. But, the method is far more general than our retarded springs with external forces. Many physical systems enjoy linearity and the principle of superposition. But, not all, for instance General Relativity and Einstein's Equations are nonlinear and as such they are far more subtle to analyze. The solution below is an **existence theorem**. It tells us what can be done, but on the other hand, it doesn't actually tell us how to accomplish the task of solving the differential equation.

Theorem 3.2.5. *unique solutions to the initial value problem for $L[y] = f$.*

Suppose $a_0, a_1, \dots, a_n, f$ are continuous on an interval I with $a_n(x) \neq 0$ for all $x \in I$. Suppose $x_0 \in I$ and $y_0, y_1, \dots, y_{n-1} \in \mathbb{R}$ then there exists a **unique** function ϕ such that:

$$(1.)\quad a_n(x)\phi^{(n)}(x) + a_{n-1}(x)\phi^{(n-1)}(x) + \cdots + a_1(x)\phi'(x) + a_0(x)\phi(x) = f(x)$$

for all $x \in I$ and

$$(2.)\quad \phi(x_0) = y_0,\ \phi'(x_0) = y_1,\ \phi''(x_0) = y_2,\ \dots,\ \phi^{(n-1)}(x_0) = y_{n-1}.$$

The linear differential equation $L[y] = f$ on an interval I paired with the n-conditions $y(x_0) = y_0,\ y'(x_0) = y_1,\ y''(x_0) = y_2,\ \dots,\ y^{(n-1)}(x_0) = y_{n-1}$ is called an **initial value problem** (IVP). The theorem above simply says that there is a unique solution to the initial value problem for any linear n-th order ODE with continuous coefficients. The proof of this theorem can be found in many advanced calculus or differential equations texts. See Chapter 13 of Nagle Saff and Snider for some discussion. We can't cover it here because we need ideas about convergence of sequences of functions. If you are interested you should return to this theorem after you have completed the real analysis course. Proof aside, we will see how this theorem works dozens if not hundreds of times as the course continues.

I'll illustrate the theorem with some examples.

Example 3.2.6. *The solution of of $y' = y$ with $y(0) = 1$ is given by $y = e^x$. Here $L[y] = y' - y$ and the coefficient functions are $a_1 = 1$ and $a_0 = -1$. These constant coefficients are continuous on $\mathbb{R}$ and $a_1 = 1 \neq 0$ on $\mathbb{R}$ as well. It follows from Theorem 3.2.5 that the unique solution with $y(0) = 1$ should exist on $\mathbb{R}$.*

Example 3.2.7. *The general solution of $y'' + y$ is given by*

$$y = c_1 \cos(x) + c_2 \sin(x)$$

by the method of reduction of order shown in Example 2.4.11. Theorem 3.2.5 indicates that there is a unique choice of c_1, c_2 to produce a particular set of initial conditions. For example: the solution of of $y'' + y = 0$ with $y(0) = 1, y'(0) = 1$ is given by $y = \cos(x) + \sin(x)$. Here $L[y] = y'' + y$ and the coefficient functions are $a_2 = 1, a_1 = 0$ and $a_0 = 1$. These constant coefficients are continuous on $\mathbb{R}$ and $a_2 = 1 \neq 0$ on $\mathbb{R}$ as well. Once more we see from Theorem 3.2.5 that the unique solution with $y(0) = 1, y'(0) = 1$ should exist on $\mathbb{R}$, and it does!

Example 3.2.8. *The general solution of $x^3y'' + xy' - y = 0$ is given by*

$$y = c_1 x + c_2 x e^{1/x}$$

Observe that $a_2(x) = x^3$ and $a_1(x) = x$ and $a_0(x) = -1$. These coefficient functions are continuous on $\mathbb{R}$, however, $a_2(0) = 0$. We can only expect, from Theorem 3.2.5, that solutions to exist on $(0, \infty)$ or $(-\infty, 0)$. This is precisely the structure of the general solution. I leave it to the reader to verify that the initial value problem has a unique solution on either $(0, \infty)$ or $(-\infty, 0)$.

Example 3.2.9. *Consider $y^{(n)}(t) = 0$. If we integrate n-times then we find (absorbing any fractions of integration into the constants for convenience)*

$$y(t) = c_1 + c_2 t + c_3 t^2 + \cdots + c_n t^{n-1}$$

is the general solution. Is there a unique solution to the initial value problem here? Theorem 3.2.5 indicates yes since $a_n = 1$ is nonzero on $\mathbb{R}$ and all the other coefficient functions are clearly continuous. Once more I leave the proof to the reader[2], but as an example $y''' = 0$ with $y(0) = 1, y'(0) = 1$ and $y''(0) = 2$ is solved uniquely by $y(t) = t^2 + t + 1$.

We see that there seem to be n-distinct functions forming the solution to an n-th order linear ODE. We need to develop some additional theory to make this idea of *distinct* a bit more precise. For example, we would like to count e^x and $2e^x$ as the same function since multiplication by 2 in our context could easily be absorbed into the constant. On the other hand, e^{2x} and e^{3x} are distinct functions.

[2] this makes a nice linear algebra problem

Definition 3.2.10. *linear independence of functions on an interval I.*

Let I be an interval of real numbers. We say the set of functions $\{f_1, f_2, f_3, \dots, f_m\}$ are **linearly independent (LI)** on I iff

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_m f_m(x) = 0$$

for all $x \in I$ implies $c_1 = c_2 = \cdots = c_m = 0$. Conversely, if $\{f_1, f_2, f_3, \dots, f_m\}$ are not linearly independent on I then they are said to be **linearly dependent** on I.

It is not hard to show that if $\{f_1, f_2, f_3, \dots, f_m\}$ is linearly dependent set of functions on I then there is at least one function, say f_j such that

$$f_j = c_1 f_1 + c_2 f_2 + \cdots + c_{j-1} f_{j-1} + c_{j+1} f_{j+1} + \cdots + c_n f_n.$$

This means that the function f_j is redundant. If these functions are solutions to $L[y] = 0$ then we don't really need f_j since the other $n-1$ functions can produce the same solutions under linear combinations. On the other hand, if the set of solutions is linearly independent then every function in the set is needed to produce the general solution. As a point of conversational convenience let us adopt the following convention: f_1 **and** f_2 **are independent on** I **iff** $\{f_1, f_2\}$ **are linearly independent on** I.

I may discuss direct application of the defintion above in lecture, however it is better to think about the construction to follow here. We seek a convenient computational characterization of linear independence of functions. Suppose that $\{y_1, y_2, y_3, \dots, y_m\}$ is linearly **independent** set of functions on I which are at least $(n-1)$-times differentiable. Furthermore, suppose for all $x \in I$

$$c_1 y_1(x) + c_2 y_2(x) + \cdots + c_m y_m(x) = 0.$$

Differentiate to obtain for all $x \in I$:

$$c_1 y_1'(x) + c_2 y_2'(x) + \cdots + c_m y_m'(x) = 0.$$

Differentiate again to obtain for all $x \in I$:

$$c_1 y_1''(x) + c_2 y_2''(x) + \cdots + c_m y_m''(x) = 0.$$

Continue differentiating until we obtain for all $x \in I$:

$$c_1 y_1^{(m-1)}(x) + c_2 y_2^{(m-1)}(x) + \cdots + c_m y_m^{(m-1)}(x) = 0.$$

Let us write these m-equations in matrix notation[3]

$$\begin{bmatrix} y_1(x) & y_2(x) & \cdots & y_m(x) \\ y_1'(x) & y_2'(x) & \cdots & y_m'(x) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(m-1)}(x) & y_2^{(m-1)}(x) & \cdots & y_m^{(m-1)}(x) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

In linear algebra we show that the linear equation $A\vec{x} = \vec{b}$ has a unique solution iff $det(A) \neq 0$. Since we have assumed linear independence of $\{y_1, y_2, y_3, \dots, y_m\}$ on I we know $c_1 = c_2 = \cdots = c_m = 0$ is the only solution of the system above for each $x \in I$. Therefore, the coefficient matrix must have nonzero determinant[4] on all of I. This determinant is called the **Wronskian**.

[3] don't worry too much if you don't know matrix math just yet, we will cover some of the most important matrix computations a little later in this course, for now just think of it as a convenient notation

[4] have no fear, I will soon remind you how we calculate determinants, you saw the pattern before with cross products in calculus III

Definition 3.2.11. *Wronskian of functions $y_1, y_2, \dots, y_m$ at x.*

$$W(y_1, y_2, \dots, y_m; x) = det \begin{bmatrix} y_1(x) & y_2(x) & \cdots & y_m(x) \\ y_1'(x) & y_2'(x) & \cdots & y_m'(x) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(m-1)}(x) & y_2^{(m-1)}(x) & \cdots & y_m^{(m-1)}(x) \end{bmatrix}.$$

It is clear from the discussion preceding this definition that we have the following proposition:

Theorem 3.2.12. *nonzero Wronskian on I implies linear independence on I.*

If $W(y_1, y_2, \dots, y_m; x) \neq 0$ for each $x \in I$ then $\{y_1, y_2, y_3, \dots, y_m\}$ is linearly independent on I.

Let us pause to introduce the formulas for the determinant of a square matrix. We define,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Then the 3×3 case is defined in terms of the 2×2 formula as follows:

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \cdot \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \cdot \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \cdot \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

and finally the 4×4 determinant is given by

$$\det \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix} = a \cdot \det \begin{pmatrix} f & g & h \\ j & k & l \\ n & o & p \end{pmatrix} - b \cdot \det \begin{pmatrix} e & g & h \\ i & k & l \\ m & o & p \end{pmatrix}$$
$$+ c \cdot \det \begin{pmatrix} e & f & h \\ i & j & l \\ m & n & p \end{pmatrix} - d \cdot \det \begin{pmatrix} e & f & g \\ i & j & k \\ m & n & o \end{pmatrix}$$

Expanding the formula for the determinant in terms of lower order determinants is known as *Laplace's expansion by minors*. It can be shown, after considerable effort, this is the same as defining the deteminant as the completely antisymmetric multilinear combination of the rows of A:

$$det(A) = \sum_{i_1, i_2, \dots, i_n} \epsilon_{i_1, i_2, \dots, i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}.$$

See my linear algebra notes if you want to learn more. For the most part we just need the 2×2 or 3×3 for examples.

Example 3.2.13. *Consider $y_1 = e^{ax}$ and $y_2 = e^{bx}$ for $a, b \in \mathbb{R}$ with $a \neq b$. The Wronskian is*

$$W(e^{ax}, e^{bx}; x) = det \begin{bmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{bmatrix} = \begin{bmatrix} e^{ax} & e^{bx} \\ ae^{ax} & be^{bx} \end{bmatrix} = e^{ax}be^{bx} - e^{bx}ae^{ax} = (b-a)e^{(a+b)x}.$$

Since $a - b \neq 0$ it follows $W(e^{ax}, e^{bx}, x) \neq 0$ on $\mathbb{R}$ and we find $\{e^{ax}, e^{bx}\}$ is LI on $\mathbb{R}$.

Example 3.2.14. *Consider $y_1(t) = 1$ and $y_2(t) = t$ and $y_3(t) = t^2$. The Wronskian is*

$$W(1, t, t^2; t) = det \begin{bmatrix} 1 & t & t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 2 \end{bmatrix} = (1)(1)(2) = 2.$$

Clearly $W(1, t, t^2; t) \neq 0$ for all $t \in \mathbb{R}$ and we find $\{1, t, t^2\}$ is LI on $\mathbb{R}$.

Example 3.2.15. *Consider* $y_1(x) = x$, $y_2(x) = \cosh(x)$, $y_3(x) = \sinh(x)$. *Calculate* $W(x, \cosh(x), \sinh(x); x) =$

$$\begin{aligned}
&= det \begin{bmatrix} x & \cosh(x) & \sinh(x) \\ 1 & \sinh(x) & \cosh(x) \\ 0 & \cosh(x) & \sinh(x) \end{bmatrix} \\
&= x det \begin{bmatrix} \sinh(x) & \cosh(x) \\ \cosh(x) & \sinh(x) \end{bmatrix} - \cosh(x) det \begin{bmatrix} 1 & \cosh(x) \\ 0 & \sinh(x) \end{bmatrix} + \sinh(x) det \begin{bmatrix} 1 & \sinh(x) \\ 0 & \cosh(x) \end{bmatrix} \\
&= x\big[\sinh^2(x) - \cosh^2(x)\big] - \cosh(x)\big[1\sinh(x) - 0\cosh(x)\big] + \sinh(x)\big[1\cosh(x) - 0\sinh(x)\big] \\
&= -x.
\end{aligned}$$

Clearly $W(x, \cosh(x), \sinh(x); x) \neq 0$ *for all* $x \neq 0$. *It follows* $\{x, \cosh(x), \sinh(x)\}$ *is LI on any interval which does not contain zero.*

The interested reader is apt to ask: is $\{x, \cosh(x), \sinh(x)\}$ linearly dependent on an interval which does contain zero? The answer is no. In fact:

Theorem 3.2.16. *Wronskian trivia.*

> Suppose $\{y_1, y_2, \dots, y_m\}$ are $(n-1)$-times differentiable on an interval I. If $\{y_1, y_2, \dots, y_m\}$ is linearly dependent on I then $W(y_1, y_2, \dots, y_m; x) = 0$ for all $x \in I$. Conversely, if there exists $x_0 \in I$ such that $W(y_1, y_2, \dots, y_m; x) \neq 0$ then $\{y_1, y_2, \dots, y_m\}$ is LI on I.

The still interested reader might ask: "what if the Wronskian is zero at all points of some interval? Does that force linear dependence?". Again, no. Here's a standard example that probably dates back to a discussion by Peano and others in the late 19-th century:

Example 3.2.17. *The functions* $y_1(x) = x^2$ *and* $y_2(x) = x|x|$ *are linearly independent on* $\mathbb{R}$. *You can see this from supposing* $c_1x^2 + c_2x|x| = 0$ *for all* $x \in \mathbb{R}$. *Take* $x = 1$ *to obtain* $c_1 + c_2 = 0$ *and take* $x = -1$ *to obtain* $c_1 - c_2 = 0$ *which solved simultaneously yield* $c_1 = c_2 = 0$. *However,*

$$W(x^2, x|x|; x) = det \begin{bmatrix} x^2 & x|x| \\ 2x & 2|x| \end{bmatrix} = 0.$$

The Wronskian is useful for testing linear-dependence of complete solution sets of a linear ODE.

Theorem 3.2.18. *Wronskian on a solution set of a linear ODE.*

> Suppose $L[y] = 0$ is an n-th order linear ODE on an interval I and $y_1, y_2, \dots, y_n$ are solutions on I. If there exists $x_0 \in I$ such that $W(y_1, y_2, \dots, y_n; x_0) \neq 0$ then $\{y_1, y_2, \dots, y_n\}$ is LI on I. On the other hand, if there exists $x_0 \in I$ such that $W(y_1, y_2, \dots, y_n; x_0) = 0$ then $\{y_1, y_2, \dots, y_n\}$ is linearly dependent on I

Notice that the number of solutions considered must match the order of the equation. It turns out the theorem does not apply to smaller sets of functions. It is possible for the Wronskian of two solutions to a third order ODE to vanish even though the functions are linearly independent. The most interesting proof of the theorem above is given by Abel's formula. I'll show how to derive it in the $n = 2$ case to begin:

Let a_0, a_1, a_2 be continuous functions on an interval I with $a_2(x) \neq 0$ for each $x \in I$. Suppose $a_2y'' + a_1y' + a_0y = 0$ has solutions y_1, y_2 on I. Consider the Wronskian $W(x) = y_1y_2' - y_2y_1'$. Something a bit interesting happens as we calculate the derivative of W,

$$W' = y_1'y_2' + y_1y_2'' - y_2'y_1' - y_2y_1'' = y_1y_2'' - y_2y_1''.$$

However, y_1 and y_2 are solutions of $a_2y'' + a_1y' + a_0y = 0$ hence

$$y_1'' = -\frac{a_1}{a_2}y_1' - \frac{a_0}{a_2}y_1 \qquad \& \qquad y_2'' = -\frac{a_1}{a_2}y_2' - \frac{a_0}{a_2}y_2$$

Therefore,

$$W' = y_1\left(-\frac{a_1}{a_2}y_2' - \frac{a_0}{a_2}y_2\right) - y_2\left(-\frac{a_1}{a_2}y_1' - \frac{a_0}{a_2}y_1\right) = -\frac{a_1}{a_2}\left(y_1y_2' - y_2y_1'\right) = -\frac{a_1}{a_2}W$$

We can solve $\frac{dW}{dx} = \frac{a_1}{a_2}W$ by separation of variables:

$$\boxed{W(x) = C\,exp\left[-\int \frac{a_1}{a_2}\,dx\right]} \qquad \Leftarrow \text{ Abel's Formula.}$$

It follows that either $C = 0$ and $W(x) = 0$ for all $x \in I$ or $C \neq 0$ and $W(x) \neq 0$ for all $x \in I$.

It is a bit surprising that Abel's formula does not involve a_0 directly. It is fascinating that this continues to be true for the n-th order problem: if $y_1, y_2, \dots, y_n$ are solutions of $a_ny^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0$ and W is the Wronskian of the given n-functions then W is gvien by Abel's formula $W(x) = C\,exp\left[\int \frac{a_{n-1}}{a_n}\,dx\right]$. You can skip the derivation that follows if you wish. What follows is an example of tensor calculus: let $Y = [y_1, y_2, \dots, y_n]$ thus $Y' = [y_1', y_2', \dots, y_n']$ and $Y^{(n-1)} = [y_1^{(n-1)}, y_2^{(n-1)}, \dots, y_n^{(n-1)}]$. The Wronskian is given by

$$W = \sum_{i_1,i_2,\dots,i_n=1}^{n} \epsilon_{i_1i_2\dots i_n} Y_{i_1}Y_{i_2}' \cdots Y_{i_n}^{(n-1)}$$

Apply the product rule for n-fold products on each summand in the above sum,

$$W' = \sum_{i_1,i_2,\dots,i_n=1}^{n} \epsilon_{i_1i_2\dots i_n}\left(Y_{i_1}'Y_{i_2}' \cdots Y_{i_n}^{(n-1)} + Y_{i_1}Y_{i_2}''Y_{i_3}'' \cdots Y_{i_n}^{(n-1)} + \cdots + Y_{i_1}Y_{i_2}' \cdots Y_{i_{n-1}}^{(n-2)}Y_{i_n}^{(n)}\right)$$

The term $Y_{i_1}'Y_{i_2}' \cdots Y_{i_n}^{(n-1)} = Y_{i_2}'Y_{i_1}' \cdots Y_{i_n}^{(n-1)}$ hence is symmetric in the pair of indices i_1, i_2. Next, the term $Y_{i_1}Y_{i_2}''Y_{i_3}'' \cdots Y_{i_n}^{(n-1)}$ is symmetric in the pair of indices i_2, i_3. This patterns continues up to the term $Y_{i_1}Y_{i_2}' \cdots Y_{i_{n-2}}^{(n-1)}Y_{i_{n-1}}^{(n-2)}Y_{i_n}^{(n-1)}$ which is symmetric in the i_{n-2}, i_{n-1} indices. In contrast, the completely antisymmetric symbol $\epsilon_{i_1i_2\dots i_n}$ is antisymmetric in each possible pair of indices. Note that if $S_{ij} = S_{ji}$ and $A_{ij} = -A_{ji}$ then

$$\sum_i\sum_j S_{ij}A_{ij} = \sum_i\sum_j -S_{ji}A_{ji} = -\sum_j\sum_i S_{ji}A_{ji} = -\sum_i\sum_j S_{ij}A_{ij} \quad \Rightarrow \quad \sum_i\sum_j S_{ij}A_{ij} = 0.$$

If we sum an antisymmetric object against a symetric object then the result is zero. It follows that only one term remains in calculation of W':

$$W' = \sum_{i_1,i_2,\dots,i_n=1}^{n} \epsilon_{i_1i_2\dots i_n} Y_{i_1}Y_{i_2}' \cdots Y_{i_{n-1}}^{(n-2)}Y_{i_n}^{(n)} \qquad (\star)$$

Recall that $y_1, y_2, \dots, y_n$ are solutions of $a_ny^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0$ hence

$$y_j^{(n)} = -\frac{a_{n-1}}{a_n}y_j^{(n-1)} - \cdots - \frac{a_1}{a_n}y_j' - \frac{a_0}{a_n}y_j = 0$$

for each $j = 1, 2, \dots, n$. But, this yields

$$Y^{(n)} = -\frac{a_{n-1}}{a_n}Y^{(n-1)} - \cdots - \frac{a_1}{a_n}Y' - \frac{a_0}{a_n}Y$$

Substitute this into $\star$,

$$\begin{aligned}
W' &= \sum_{i_1,i_2,\dots,i_n=1}^{n} \epsilon_{i_1 i_2 \dots i_n} Y_{i_1} Y'_{i_2} \cdots Y^{(n-2)}_{i_{n-1}} \left[-\frac{a_{n-1}}{a_n} Y^{(n-1)} - \cdots - \frac{a_1}{a_n} Y' - \frac{a_0}{a_n} Y \right]_{i_n} \\
&= \sum_{i_1,i_2,\dots,i_n=1}^{n} \epsilon_{i_1 i_2 \dots i_n} \left(-\frac{a_{n-1}}{a_n} Y_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_n} - \cdots - \frac{a_1}{a_n} Y_{i_1} Y'_{i_2} \cdots Y'_{i_n} - \frac{a_0}{a_n} Y_{i_1} Y'_{i_2} \cdots Y_{i_n} \right) \\
&= -\frac{a_{n-1}}{a_n} \sum_{i_1,i_2,\dots,i_n=1}^{n} \epsilon_{i_1 i_2 \dots i_n} Y_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_n} \qquad \star\star \\
&= -\frac{a_{n-1}}{a_n} W.
\end{aligned}$$

The $\star\star$ step is based on the observation that the index pairs i_1, i_n and i_2, i_n etc... are symmetric in the line above it hence as they are summed against the completely antisymmetric symbol those terms vanish. Alternatively, and equivalently, you could apply the multilinearity of the determinant paired with the fact that a determinant with any two repeated rows vanishes. Linear algebra aside, we find $W' = -\frac{a_{n-1}}{a_n} W$ thus Abel's formula $W(x) = C\,exp\big[\int \frac{a_{n-1}}{a_n}\,dx\big]$ follows immediately.

Solution sets of functions reside in function space. As a vector space, function space is infinite dimensional. The matrix techniques you learn in the linear algebra course do not apply to the totallity of function space. Appreciate the Wronskian says what it says. In any event, we should continue our study of DEqns at this point since we have all the tools we need to understand LI in this course.

Definition 3.2.19. *fundamental solutions sets of linear ODEs.*

> Suppose $L[y] = f$ is an n-th order linear differential equation on an interval I. We say $S = \{y_1, y_2, \dots, y_n\}$ is a **fundamental solution set** of $L[y] = f$ iff S is linearly independent set of solutions to the homogeneous equation; $L[y_j] = 0$ for $j = 1, 2, \dots n$.

Example 3.2.20. *The differential equation $y'' + y = f$ has fundamental solution set $\{\cos(x), \sin(x)\}$. You can easily verify that $W(\cos(x), \sin(x); x) = 1$ hence linear independence is established the given functions. Moreover, it is simple to check $y'' + y = 0$ has sine and cosine as solutions. The formula for f is irrelevant to the fundamental solution set. Generally, the fundamental solution set is determined by the structure of L when we consider the general problem $L[y] = f$.*

Theorem 3.2.21. *existence of a fundmental solution set.*

> If $L[y] = f$ is an n-th order linear differential equation with continuous coefficient functions on an interval I then there exists a fundamental solution set $S = \{y_1, y_2, \dots, y_n\}$ on I.

Proof: Theorem 3.2.5 applies. Pick $x_0 \in I$ and use the existence theorem to obtain the solution y_1 subject to

$$y_1(x_0) = 1, \quad y_1'(x_0) = 0, \quad y_1''(x_0) = 0, \ \dots \ , y_1^{(n-1)}(x_0) = 0.$$

Apply the theorem once more to select solution y_2 with:

$$y_2(x_0) = 0, \quad y_2'(x_0) = 1, \quad y_2''(x_0) = 0, \ \dots \ , y_2^{(n-1)}(x_0) = 0.$$

Then continue in this fashion selecting solutions $y_3, y_4, \dots, y_{n-1}$ and finally y_n subject to

$$y_n(x_0) = 0, \quad y_n'(x_0) = 0, \quad y_n''(x_0) = 0, \ \dots \ , y_n^{(n-1)}(x_0) = 1.$$

It remains to show that the solution set $\{y_1, y_2, \dots, y_n\}$ is indeed linearly independent on I. Calculate the Wronskian at $x = x_0$ for the solution set $\{y_1, y_2, \dots, y_n\}$: abbreviate $W(y_1, y_2, \dots, y_n; x)$ by $W(x)$ for the remainder of this proof:

$$W(x_0) = det \begin{bmatrix} y_1(x_0) & y_2(x_0) & \cdots & y_m(x_0) \\ y_1'(x_0) & y_2'(x_0) & \cdots & y_m'(x_0) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(n-1)}(x_0) & y_2^{(n-1)}(x_0) & \cdots & y_n^{(n-1)}(x_0) \end{bmatrix} = det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = 1.$$

Therefore, by Abel's formula, the Wronskian is nonzero on the whole interval I and it follows the solution set is LI. □

Theorem 3.2.22. *general solution of the homogeneous linear n-th order problem.*

If $L[y] = f$ is an n-th order linear differential equation with continuous coefficient functions on an interval I with fundamental solution set $S = \{y_1, y_2, \dots, y_n\}$ on I. Then any solution of $L[y] = 0$ can be expressed as a linear combination of the fundamental solution set: that is, there exist constants $c_1, c_2, \dots, c_n$ such that:

$$y = c_1y_1 + c_2y_2 + \cdots + c_ny_n.$$

Proof: Suppose $S = \{y_1, y_2, \dots, y_n\}$ is a fundamental solution set of $L[y] = 0$ on I. Furthermore, suppose y is a solution; $L[y] = 0$. We seek to find $c_1, c_2, \dots, c_n$ such that $y = c_1y_1 + c_2y_2 + \cdots + c_ny_n$. Consider a particular point $x_0 \in I$, we need that solution y and its derivatives $(y', y'', ...)$ up to order $(n-1)$ match with the proposed linear combination of the solution set:

$$y(x_0) = c_1y_1(x_0) + c_2y_2(x_0) + \cdots + c_ny_n(x_0).$$

$$y'(x_0) = c_1y_1'(x_0) + c_2y_2'(x_0) + \cdots + c_ny_n'(x_0).$$

continuing, up to the $(n-1)$-th derivative

$$y^{(n-1)}(x_0) = c_1y_1^{(n-1)}(x_0) + c_2y_2^{(n-1)}(x_0) + \cdots + c_ny_n^{(n-1)}(x_0).$$

It is instructive to write this as a matrix problem:

$$\begin{bmatrix} y(x_0) \\ y'(x_0) \\ \vdots \\ y^{(n-1)}(x_0) \end{bmatrix} = \begin{bmatrix} y_1(x_0) & y_2(x_0) & \cdots & y_n(x_0) \\ y_1'(x_0) & y_2'(x_0) & \cdots & y_n'(x_0) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(n-1)}(x_0) & y_2^{(n-1)}(x_0) & \cdots & y_n^{(n-1)}(x_0) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

The coefficient matrix has nonzero determinant (it is the Wronskian at $x = x_0$) hence this system of equations has a unique solution. Therefore, we can select constants $c_1, c_2, \dots, c_n$ such that the solution $y = c_1y_1 + c_2y_2 + \cdots + c_ny_n$. □

In fact, the proof shows that these constants are unique for a given fundamental solution set. Each solution is uniquely specified by the constants $c_1, c_2, \dots, c_n$. When I think about the solution of a linear ODE, I always think of the constants in the general solution as the reflection of the reality that a given DEqn can be assigned many different initial conditions. However, once the initial condition is given the solution is specified uniquely.

Finally we turn to the nonhomogeneous problem. I present the theory here, however, the computational schemes are given much later in this chapter.

Theorem 3.2.23. *general solution of the nonhomogeneous linear n-th order problem.*

If $L[y] = f$ is an n-th order linear differential equation with continuous coefficient functions on an interval I with fundamental solution set $S = \{y_1, y_2, \dots, y_n\}$ on I. Then any solution of $L[y] = f$ can be expressed as a linear combination of the fundamental solution set and a function y_p with $L[y_p] = f$ known as the **particular solution** :

$$y = c_1y_1 + c_2y_2 + \cdots + c_ny_n + y_p.$$

Proof: Theorem 3.2.5 applies, it follows there are many solutions to $L[y] = f$, one for each set of initial conditions. Suppose y and y_p are two solutions to $L[y] = f$. Observe that

$$L[y - y_p] = L[y] - L[y_p] = f - f = 0.$$

Therefore, $y_h = y - y_p$ is a solution of the homogeneous ODE $L[y] = 0$ thus Theorem 3.2.22 we can write y_h as a linear comination of the fundamental solutions: $y_h = c_1y_1 + c_2y_2 + \cdots + c_ny_n$. But, $y = y_p + y_h$ and the theorem follows. □

Example 3.2.24. *Suppose $L[y] = f$ is a second order linear ODE and $y = e^x + x^2$ and $z = \cos(x) + x^2$ are solutions. Then*

$$L[y - z] = L[y] - L[z] = f - f = 0$$

hence $y - z = (e^x + x^2) - (\cos(x) + x^2) = e^x - \cos(x)$ gives a homogeneous solution $y_1(x) = e^x - \cos(x)$. Notice that $w = y + 2y_1 = e^x + x^2 + 2(e^x - \cos(x)) = 3e^x - 2\cos(x) + x^2$ is also a solution since $L[y + 2y_1] = L[y] + 2L[y_1] = f + 0 = f$.

The example above is important because it illustrates that we can extract homogeneous solutions from particular solutions. Physically speaking, perhaps you might be faced with the same system subject to several different forces. If solutions are observed for $L[y] = F_1$ and $L[y] = 2F_1$ then we can deduce the general solution set of $L[y] = 0$. In particular, this means you could deduce the mass and spring constant of a particular spring-mass system by observing how it responds to a pair of forces. More can be said here, we'll return to these thoughts as we later discuss the *principle of superposition*.

3.3 constant coefficient homogeneous problem

In Section 3.1 we established many facts about operators. Let us review those concepts briefly once more, this time without proof. Let $L = P(D)$ for some polynomial with real coefficients $P(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$. By the fundamental theorem of algebra we can factor P into n-linear factors. In particular, if $P(x) = 0$ has solutions $r_1, r_2, \dots, r_k$ then the factor theorem implies that there are real constants $m_1, m_2, \dots, m_k$ with $m_1 + m_2 + \cdots + m_k = n$ and

$$P(x) = a_n(x - r_1)^{m_1}(x - r_2)^{m_2} \cdots (x - r_k)^{m_k}$$

I include the possibility that r_j could be complex. P is a polynomial with real coefficients, it follows that if r_j is a complex zero then the complex conjugate r_j^* also has $P(r_j^*) = 0$. By Theorem 3.2.4 the polynomial of the differentiation operator $P(D)$ shares the same factorization:

$$L = P(D) = a_n(D - r_1)^{m_1}(D - r_2)^{m_2} \cdots (D - r_k)^{m_k}$$

We wish to solve the differential equation $P(D)[y] = 0$. The following facts hold for both real and complex zeros. However, understand that when r_j is complex the corresponding solutions are likewise complex:

(1.) if $D = d/dt$ then the DEqn $(D - r)[y] = 0$ has solution $y = e^{rt}$.

(2.) if $D = d/dt$ then the DEqn $(D - r)^2[y] = 0$ has two solutions $y = e^{rt}, te^{rt}$.

(3.) if $D = d/dt$ then the DEqn $(D-r)^3[y] = 0$ has three solutions $y = e^{rt}, te^{rt}, t^2e^{rt}$.

(4.) if $D = d/dt$ then the DEqn $(D-r)^m[y] = 0$ has m-solutions $y = e^{rt}, te^{rt}, \dots, t^{m-1}e^{rt}$

(5.) $\{e^{rt}, te^{rt}, \dots, t^{m-1}e^{rt}\}$ is a LI set of functions (on $\mathbb{R}$ or $\mathbb{C}$).

Let us unravel the complex case into real notation. Suppose $r = \alpha + i\beta$ then $r^* = \alpha - i\beta$. Note:

$$e^{rt} = e^{\alpha t + i\beta t} = e^{\alpha t}\cos(\beta t) + ie^{\alpha t}\sin(\beta t)$$

$$e^{r^*t} = e^{\alpha t - i\beta t} = e^{\alpha t}\cos(\beta t) - ie^{\alpha t}\sin(\beta t)$$

Observe that the both complex functions give the same real solution set:

$$Re(e^{\alpha t \pm i\beta t}) = e^{\alpha t}\cos(\beta t) \qquad \& \qquad Im(e^{\alpha t \pm i\beta t}) = \pm e^{\alpha t}\sin(\beta t)$$

If $(D-r)^m[y] = 0$ has m-**complex** solutions $y = e^{rt}, te^{rt}, \dots, t^{m-1}e^{rt}$ then $(D-r)^m[y] = 0$ possesses the $2m$-**real solutions**

$$e^{\alpha t}\cos(\beta t),\ e^{\alpha t}\sin(\beta t),\ te^{\alpha t}\cos(\beta t),\ te^{\alpha t}\sin(\beta t),\ \dots,\ t^{m-1}e^{\alpha t}\cos(\beta t),\ t^{m-1}e^{\alpha t}\sin(\beta t).$$

It should be clear how to assemble the general solution to the general constant coefficient problem $P(D)[y] = 0$. I will abstain from that notational quagmire and instead illustrate with a series of examples.

Example 3.3.1. Problem: *Solve* $y'' + 3y' + 2y = 0$.
Solution: *Note the differential equation is* $(D^2+3D+2)[y]$. *Hence* $(D+1)(D+2)[y] = 0$. *We find solutions* $y_1 = e^{-x}$ *and* $y_2 = e^{-2x}$ *therefore the general solution is* $\boxed{y = c_1e^{-x} + c_2e^{-2x}}$.

Example 3.3.2. Problem: *Solve* $y'' - 3y' + 2y = 0$.
Solution: *Note the differential equation is* $(D^2-3D+2)[y]$. *Hence* $(D-1)(D-2)[y] = 0$. *We find solutions* $y_1 = e^{x}$ *and* $y_2 = e^{2x}$ *therefore the general solution is* $\boxed{y = c_1e^{x} + c_2e^{2x}.}$

Example 3.3.3. Problem: *Solve* $y^{(4)} - 5y'' + 4y = 0$.
Solution: *Note the differential equation is* $(D^4 - 5D + 4)[y]$. *Note that*

$$D^4 - 5D + 4 = (D^2-1)(D^2-4) = (D+1)(D-1)(D+2)(D-2)$$

It follows that the differential equation factors to $(D+1)(D+2)(D-1)(D-2)[y] = 0$ *and the general solution reads*

$$\boxed{y = c_1e^{-x} + c_2e^{-2x} + c_3e^{x} + c_4e^{2x}.}$$

You should notice that I do **not** state that $D = \pm 1$ or $D = \pm 2$ in the example above. Those equations are illogical nonsense. I am using the theory we've developed in this chapter to extract solutions from inspection of the factored form. If you really want to think in terms of roots instead of factors then I would advise that you use the following fact:

$$\boxed{P(D)[e^{\lambda t}] = P(\lambda)e^{\lambda t}.}$$

I exploited this identity to solve the second order problem in our first lecture on the n-th order problem. Solutions to $P(\lambda) = 0$ are called the **characteristic values** of the DEqn $P(D)[y] = 0$. The equation $P(\lambda) = 0$ is called the **characteristic equation**.

Example 3.3.4. Problem: *Solve* $y^{(4)} - 5y'' + 4y = 0$.
Solution: *Let* $P(D) = D^4 - 5D + 4$ *thus the DEqn is* $P(D)[y] = 0$. *Note that* $P(\lambda) = \lambda^4 - 5\lambda + 4$.

$$\lambda^4 - 5\lambda + 4 = (\lambda^2-1)(\lambda^2-4) = (\lambda+1)(\lambda-1)(\lambda+2)(\lambda-2)$$

Hence, the solutions of $P(\lambda) = 0$ *are* $\lambda_1 = -1, \lambda_1 = -2, \lambda_3 = 1$ *and* $\lambda_4 = 2$ *the* **characteristic values** *of* $P(D)[y] = 0$. *The general solution follows:*

$$\boxed{y = c_1e^{-x} + c_2e^{-2x} + c_3e^{x} + c_4e^{2x}.}$$

We can also group the exponential functions via the hyperbolic sine and cosine. Since

$$\cosh(x) = \frac{1}{2}(e^x + e^{-x}) \qquad \& \qquad \sinh(x) = \frac{1}{2}(e^x - e^{-x})$$

we have $e^x = \cosh(x) + \sinh(x)$ and $e^{-x} = \cosh(x) - \sinh(x)$. Thus,

$$c_1 e^{-x} + c_3 e^x = c_1(\cosh(x) - \sinh(x)) + c_3(\cosh(x) + \sinh(x)) = (c_1 + c_3)\cosh(x) + (c_1 - c_3)\sinh(x).$$

For a given problem we can either use exponentials or hyperbolic sine and cosine.

Example 3.3.5. **Problem:** *Solve $y'' - y = 0$ with $y(0) = 1$ and $y'(0) = 2$.*
Solution: *we find $\lambda^2 - 1 = 0$. Hence $\lambda = \pm 1$. We find general solution $y = c_1\cosh(x) + c_2\sinh(x)$ in view of the comments just above this example (worth remembering for later btw). Observe:*

$$y' = c_1\sinh(x) + c_2\cosh(x)$$

Consequently, $y(0) = c_1 = 1$ and $y'(0) = c_2 = 2$ and we find $\boxed{y = \cosh(x) + 2\sinh(x).}$

Believe it or not, the hyperbolic sine and cosine are easier to work with when we encounter this type of ODE in our study of boundary value problems in partial differential equations towards the conclusion of this course.

Example 3.3.6. **Problem:** *Solve $y^{(4)} + 2y'' + y = 0$.*
Solution: *the characteristic equation is $\lambda^4 + 2\lambda^2 + 1 = 0$. Hence $(\lambda^2 + 1)^2 = 0$. It follows that we have $\lambda = \pm i$ repeated. The general solution is found from the real and imaginary parts of e^{it} and te^{it}. Since $e^{it} = \cos(t) + i\sin(t)$ we find:*

$$\boxed{y = c_1\cos(t) + c_2\sin(t) + c_3 t\cos(t) + c_4 t\sin(t).}$$

Up to this point I have given examples where we had to factor the operator (or characteristic eqn.) to extract the solution. Sometimes we find problems where the operators are already factored. I consider a few such problems now.

Example 3.3.7. **Problem:** *Solve $(D^2 + 9)(D - 2)^3[y] = 0$ with $D = d/dx$ for a change.*
Solution: *I read from the expression above that we have $\lambda = \pm 3i$ and $\lambda = 2$ thrice. Hence,*

$$\boxed{y = c_1\cos(3x) + c_2\sin(3x) + c_3 e^{2x} + c_4 x e^{2x} + c_5 x^2 e^{2x}.}$$

Example 3.3.8. **Problem:** *Solve $(D^2 + 4D + 5)[y] = 0$ with $D = d/dx$.*
Solution: *Complete the square to see that $P(D)$ is not reducible; $D^2 + 4D + 5 = (D + 2)^2 + 1$ it follows that the characteristic values are $\lambda = -2 \pm i$ and the general solution is given from the real and imaginary parts of $e^{-2x+ix} = e^{-2x}e^{ix}$*

$$\boxed{y = c_1 e^{-2x}\cos(x) + c_2 e^{-2x}\sin(x).}$$

Example 3.3.9. **Problem:** *Solve $(D^2 + 4D - 5)[y] = 0$ with $D = d/dx$.*
Solution: *Complete the square; $D^2 + 4D - 5 = (D + 2)^2 - 9$ it follows that the characteristic values are $\lambda = -2 \pm 3$ or $\lambda_1 = 1$ or $\lambda_2 = -5$*

$$\boxed{y = c_1 e^x + c_2 e^{-5x}.}$$

Of course, if you love hyperbolic sine and cosine then perhaps you would prefer that we see from $((D+2)^2 - 9)[y] = 0$ the solutions

$$y = b_1 e^{-2x}\cosh(3x) + b_2 e^{-2x}\sinh(3x)$$

as the natural expression of the general solution. In invite the reader to verify the solution above is just another way to write the solution $y = c_1 e^x + c_2 e^{-5x}$.

Example 3.3.10. **Problem:** *Solve* $(D^2+6D+15)(D^2+1)(D^2-4)[y]=0$ *with* $D=d/dx$.
Solution: *Completing the square gives* $((D+3)^2+6)(D^2+1)(D^2-4)[y]=0$ *hence we find characteristic values of* $\lambda=-3\pm i\sqrt{6}, \pm i, \pm 2$. *The general solution follows:*

$$\boxed{y=c_1e^{-3x}\cos(\sqrt{6}x)+c_2e^{-3x}\sin(\sqrt{6}x)+c_3\cos(x)+c_4\sin(x)+c_5e^{2x}+c_6e^{-2x}.}$$

The example that follows is a bit more challenging since it involves both theory and a command of polynomial algebra.

Example 3.3.11. **Problem:** *Solve* $(D^5-8D^2-4D^3+32)[y]=0$ *given that* $y=\cosh(2t)$ *is a solution.*
Solution: *Straightforward factoring of the polynomial is challenging here, but I gave an olive branch. Note that if* $y=\cosh(2t)$ *is a solution then* $y=\sinh(2t)$ *is also a solution. It follows that* $(D^2-4)=(D-2)(D+2)$ *is a factor of* $D^5-8D^2-4D^3+32$. *For clarity of thought lets work on* $x^5-8x^2-4x^3+32$ *and try to factor out* x^2-4. *Long division is a nice tool for this problem. Recall:*

$$\begin{array}{r|l} & x^3 \qquad\qquad -8 \\ \hline x^2-4 & x^5-4x^3-8x^2+32 \\ & -x^5+4x^3 \\ \hline & \qquad\quad -8x^2+32 \\ & \qquad\qquad 8x^2-32 \\ \hline & \qquad\qquad\qquad\quad 0 \end{array}$$

Thus,

$$x^5-8x^2-4x^3+32=(x^2-4)(x^3-8)$$

Clearly $x^3-8=0$ *has solution* $x=2$ *hence we can factor* $(x-2)$. *I'll use long-division once more (of course, some of you might prefer synthetic division and/or have this memorized already... good)*

$$\begin{array}{r|l} & x^2+2x+4 \\ \hline x-2 & x^3 \qquad\qquad -8 \\ & -x^3+2x^2 \\ \hline & \qquad 2x^2 \\ & \quad -2x^2+4x \\ \hline & \qquad\qquad 4x-8 \\ & \qquad\quad -4x+8 \\ \hline & \qquad\qquad\qquad 0 \end{array}$$

Consequently, $x^5-8x^2-4x^3+32=(x^2-4)(x-2)(x^2+2x+4)$. *It follows that*

$$(D^5-8D^2-4D^3+32)[y]=0 \Rightarrow (D-2)^2(D+2)((D+1)^2+3)[y]=0$$

Which suggests the solution below:

$$\boxed{y=c_1e^{2t}+c_2te^{2t}+c_3e^{-2t}+c_4e^{-t}\cos(\sqrt{3}t)+c_5e^{-t}\sin(\sqrt{3}t).}$$

3.4 annihilator method for nonhomogeneous problems

In the previous section we learned how to solve **any** constant coefficient n-th order ODE. We now seek to extend the technique to the nonhomogeneous problem. Our goal is to solve:

$$L[y] = f$$

where $L = P(D)$ as in the previous section, it is a polynomial in the differentiation operator D. Suppose we find a differential operator A such that $A[f] = 0$. This is called an **annihilator** for f. Operate on $L[y] = f$ to obtain $AL[y] = A[f] = 0$. Therefore, if we have an annihilator for the forcing function f then the differential equation yields a corresponding homogeneous differential equation $AL[y] = 0$. Suppose $y = y_h + y_p$ is the general solution as discussed for Theorem 3.2.23 we have $L[y_h] = 0$ and $L[y_p] = f$. Observe:

$$AL[y_h + y_p] = A[L[y_h + y_p]] = A[f] = 0$$

It follows that the general solution to $AL[y] = 0$ will include the general solution of $L[y] = f$. The method we justify and implement in this section is commonly called the **method of undetermined coefficients**. The annihilator method shows us how to set-up the coefficients. To begin, we should work on finding annihilators to a few simple functions.

Example 3.4.1. Problem: *find an annihilator for e^x.*
Solution: *recall that e^x arises as the solution of $(D-1)[y] = 0$ therefore a natural choice for the annihilator is $A = D - 1$. This choice is* **minimal***. Observe that $A_2 = Q(D)(D - 1)$ is also an annihilator of e^x since $A_2[e^x] = Q(D)[(D - 1)[e^x]] = Q(D)[0] = 0$. There are many choices, however, we prefer the minimal annihilator. It will go without saying that all the choices that follow from here on out are minimal.*

Example 3.4.2. Problem: *find an annihilator for xe^{3x}.*
Solution: *recall that xe^{3x} arises as a solution of $(D - 3)^2[y] = 0$ hence choose $A = (D - 3)^2$.*

Example 3.4.3. Problem: *find an annihilator for $e^{3x}\cos(x)$.*
Solution: *recall that $e^{3x}\cos(x)$ arises as a solution of $((D-3)^2+1)[y] = 0$ hence choose $A = ((D-3)^2+1)$.*

Example 3.4.4. Problem: *find an annihilator for $x^2e^{3x}\cos(x)$.*
Solution: *recall that $x^2e^{3x}\cos(x)$ arises as a solution of $((D - 3)^2 + 1)^3[y] = 0$ hence choose $A = ((D - 3)^2 + 1)^3$.*

Example 3.4.5. Problem: *find an annihilator for $2e^x\cosh(2x)$.*
Solution: *observe that $2e^x\cosh(2x) = e^x(e^{2x} + e^{-2x}) = e^{3x} + e^{-x}$ and note that $(D - 3)[e^{3x}] = 0$ and $(D + 1)[e^{-x}] = 0$ thus $A = (D - 3)(D + 1)$ will do nicely.*

For those who love symmetric calculational schemes, you could also view $2e^x\cosh(2x)$ as the solution arising from $((D - 1)^2 - 4)[y] - 0$. Naturally $(D - 1)^2 - 4 = D^2 - 2D - 3 = (D - 3)(D + 1)$.

Example 3.4.6. Problem: *find an annihilator for $x^2 + e^{3x}\cos(x)$.*
Solution: *recall that $e^{3x}\cos(x)$ arises as a solution of $((D-3)^2+1)[y] = 0$ hence choose $A_1 = ((D-3)^2+1)$. Next notice that x^2 arises as a solution of $D^3[y] = 0$ hence we choose $A_2 = D^3$. Construct $A = A_1A_2$ and notice how this works: (use $A_1A_2 = A_2A_1$ which is true since these are constant coefficient operators)*

$$\begin{aligned} A_1[A_2[x^2 + e^{3x}\cos(x)]] &= A_1[A_2[x^2]] + A_2[A_1[e^{3x}\cos(x)]] \\ &= A_1[0] + A_2[0] \\ &= 0 \end{aligned}$$

because $A_1A_2 = A_2A_1$ for these constant coefficient operators. To summarize, we find $A = D^3((D - 3)^2 + 1)$ is an annihilator for $x^2 + e^{3x}\cos(x)$.

I hope you see the idea generally. If we are given a function which arises as the solution of a constant coefficient differential equation then we can use the equation to write the annihilator. You might wonder if there are other ways to find annihilators.... well, surely there are, but not usually for this course. I think the examples thus far give us a good grasp of how to kill the forcing function. Let's complete the method. We proceed by example.

Example 3.4.7. **Problem:** *find the general solution of* $y'' + y = 2e^x$
Solution: *observe* $L = D^2 + 1$ *and we face* $(D^2 + 1)[y] = 2e^x$. *Let* $A = D - 1$ *and operate on the given nonhomogeneous ODE,*

$$(D - 1)(D^2 + 1)[y] = (D - 1)[2e^x] = 0$$

We find general solution $y = c_1 e^x + c_2 \cos(x) + c_3 \sin(x)$. *Notice this is* **not** *the finished product. We should only have two constants in the general solution of this second order problem. But, remember, we insist that* $L[y] = f$ *in addition to the condition* $AL[y] = 0$ *hence:*

$$L[c_1 e^x + c_2 \cos(x) + c_3 \sin(x)] = 2e^x$$

which simplifes to $L[c_1 e^x] = 2e^x$ *since the functions* $\cos(x), \sin(x)$ *are solutions of* $L[y] = 0$. *Expanding* $L[c_1 e^x] = 2e^x$ *in detail gives us:*

$$D^2[c_1 e^x] + c_1 e^x = 2e^x \quad \Rightarrow \quad 2c_1 e^x = 2e^x \quad \Rightarrow \quad c_1 = 1.$$

Therefore we find, $\boxed{y = e^x + c_2 \cos(x) + c_3 \sin(x).}$

The notation used in the example above is not optimal for calculation. Usually I skip some of those steps because they're not needed once we understand the method. For example, once I write $y = c_1 e^x + c_2 \cos(x) + c_3 \sin(x)$ then I usually look to see which functions are in the fundamental solution set. Since $\{\cos(x), \sin(x)\}$ is a natural fundamental solution set this tells me that only the remaining function e^x is needed to construct the particular solution. Since c_1 is annoying to do algebra on, I instead use notation $y_p = Ae^x$. Next, calculate $y_p' = Ae^x$ and $y_p'' = Ae^x$ and plug these into the given ODE:

$$Ae^x + Ae^x = 2e^x \quad \Rightarrow \quad 2Ae^x = 2e^x \quad \Rightarrow \quad A = 1.$$

which brings us to the fact that $y_p = e^x$ and naturally $y_h = c_1 \cos(x) + c_2 \sin(x)$. The general solution is $y = y_h + y_p = c_1 \cos(x) + c_2 \sin(x) + e^x$.

Example 3.4.8. **Problem:** *find the general solution of* $y'' + 3y' + 2y = x^2 - 1$
Solution: *in operator notation the DEqn is* $(D^2 + 3D + 2)[y] = (D + 1)(D + 2)[y] = x^2 - 1$. *Let* $A = D^3$ *and operate on the given nonhomogeneous ODE,*

$$D^3(D + 1)(D + 2)[y] = D^3[x^2 - 1] = 0$$

The homogeneous ODE above has solutions $1, x, x^2, e^{-x}, e^{-2x}$. *Clearly the last two of these form the homogeneous solution whereas the particular solution is of the form* $y_p = Ax^2 + Bx + C$. *Calculate:*

$$y_p' = 2Ax + B, \qquad y_p'' = 2A$$

Plug this into the DEqn $y_p'' + 3y_p' + 2y_p = x^2 - 1$,

$$2A + 3(2Ax + B) + 2(Ax^2 + Bx + C) = x^2 - 1$$

multiply it out and collect terms:

$$2A + 6Ax + 3B + 2Ax^2 + 2Bx + 2C = x^2 - 1 \quad \Rightarrow \quad 2Ax^2 + (6A + 2B)x + 2A + 3B + 2C = x^2 - 1$$

this sort of equation is actually really easy to solve. We have two polynomials. When are they equal? Simple. When the coefficients match, thus calculate:

$$2A = 1, \quad 6A + 2B = 0, \quad 2A + 3B + 2C = -1$$

Clearly $A = 1/2$ hence $B = -3A = -3/2$. Solve for $C = -1/2 - A - 3B/2 = -1/2 - 1/2 + 9/4 = 5/4$. Therefore, the general solution is given by:

$$\boxed{y = c_1 e^{-x} + c_2 e^{-2x} + \frac{1}{2}x^2 - \frac{3}{2}x + \frac{5}{4}.}$$

At this point you might start to get the wrong impression. It might appear to you that the form of y_p has nothing to do with the form of y_h. That is a fortunate feature of the examples we have thus far considered. The next example features what I usually call **overlap**.

Example 3.4.9. Problem: *find the general solution of $y' - y = e^x + x$*
Solution: *Observe the annihilator is $A = (D-1)D^2$ and when we operate on $(D-1)[y] = e^x + x$ we obtain*

$$(D-1)D^2(D-1)[y] = (D-1)D^2[e^x + x] = 0 \quad \Rightarrow \quad (D-1)^2 D^2[y] = 0$$

Thus, $e^x, xe^x, 1, x$ are solutions. We find $y_h = c_1 e^x$ whereas the particular solution is of the form $y_p = Axe^x + Bx + C$. Calculate $y_p' = A(e^x + xe^x) + B$ and substitute into the DEqn to obtain:

$$A(e^x + xe^x) + B - (Axe^x + Bx + C) = e^x + x \quad \Rightarrow \quad (A - A)xe^x + Ae^x - Bx - C = e^x + x$$

We find from **equating coefficients** *of the linearly independent functions $e^x, 1, x$ that $A = 1$ and $-B = 1$ and $-C = 0$. Therefore,* $\boxed{y = c_1 e^x + xe^x - x.}$

If you look in my linear algebra notes I give a proof which shows we can equate coefficients for linearly independent sets. Usually in the calculation of y_p we find it useful to use the technique of equating coefficients to fix the **undetermined constants** A, B, C, etc...

Example 3.4.10. Problem: *find the general solution of $y'' + y = 4\cos(t)$*
Solution: *Observe the annihilator is $A = D^2 + 1$ and when we operate on $(D^2 + 1)[y] = 4\cos(t)$ we obtain*

$$(D^2 + 1)^2[y] = 0$$

Thus, $\cos(t), \sin(t), t\cos(t), t\sin(t)$ are solutions. We find $y_h = c_1 \cos(t) + c_2 \sin(t)$ whereas the particular solution is of the form $y_p = At\cos(t) + Bt\sin(t) = t(A\cos(t) + B\sin(t))$. Calculate

$$y_p' = A\cos(t) + B\sin(t) + t(-A\sin(t) + B\cos(t)) = (A + Bt)\cos(t) + (B - At)\sin(t)$$

$$y_p'' = B\cos(t) - A\sin(t) - (A + Bt)\sin(t) + (B - At)\cos(t) = (2B - At)\cos(t) - (2A + Bt)\sin(t)$$

It is nice to notice that $y_p'' = 2B\cos(t) - 2A\sin(t) - y_p$ hence $y_p'' + y_p = 4\cos(t)$ yields:

$$2B\cos(t) - 2A\sin(t) = 4\cos(t)$$

thus $2B = 4$, $-2A = 0$. Consequently, $A = 0, B = 2$ and the general solution is found to be:

$$\boxed{y = c_1 \cos(t) + c_2 \sin(t) + 2t\sin(t).}$$

From this point forward I omit the details of the annihilator method and simply propose the correct template for y_p.

Example 3.4.11. Problem: *find the general solution of $y' + y = x$*
Solution: *Observe $y_h = c_1 e^{-x}$ for the given DEqn. Let $y_p = Ax + B$ then $y_p' + y_p = A + Ax + B = x$ implies $A + B = 0$ and $A = 1$ hence $B = -1$ and we find* $\boxed{y = c_1 e^{-x} + x - 1}$.

Example 3.4.12. **Problem:** *find the general solution of* $y'' + 4y' = x$
Solution: *Observe* $\lambda^2 + 4\lambda = 0$ *gives solutions* $\lambda = 0, -4$ *hence* $y_h = c_1 + c_2 e^{-4x}$ *for the given DEqn. Let*[5] $y_p = Ax^2 + Bx$ *then* $y_p' = 2Ax + B$ *and* $y_p'' = 2A$ *hence* $y_p'' + 4y_p' = x$ *yields* $2A + 4(2Ax + B) = x$ *hence* $8Ax + 2A + 4B = x$. *Equate coefficients of* x *and* 1 *to find* $8A = 1$ *and* $2A + 4B = 0$ *hence* $A = 1/8$ *and* $B = -1/16$. *We find*

$$\boxed{y = c_1 + c_2 e^{-4x} + \frac{1}{8}x^2 - \frac{1}{16}x}.$$

Example 3.4.13. **Problem:** *find the general solution of* $y'' + 4y' = \cos(x) + 3\sin(x) + 1$
Solution: *Observe* $\lambda^2 + 4\lambda = 0$ *gives solutions* $\lambda = 0, -4$ *hence* $y_h = c_1 + c_2 e^{-4x}$ *for the given DEqn. Let* $y_p = A\cos(x) + B\sin(x) + Cx$ *then* $y_p' = -A\sin(x) + B\cos(x) + C$ *and* $y_p'' = -A\cos(x) - B\sin(x)$ *hence* $y_p'' + 4y_p' = \cos(x) + 3\sin(x) + 1$ *yields*

$$-A\cos(x) - B\sin(x) + 4(-A\sin(x) + B\cos(x) + C) = \cos(x) + 3\sin(x) + 1$$

Collecting like terms:

$$\Rightarrow \quad (4B - A)\cos(x) + (-4A - B)\sin(x) + 4C = \cos(x) + 3\sin(x) + 1$$

Equate coefficients of $\cos(x), \sin(x), 1$ *to obtain:*

$$4B - A = 1, \quad -4A - B = 3, \quad 4C = 1$$

Observe $B = -4A - 3$ *hence* $4(-4A - 3) - A = 1$ *or* $-17A - 12 = 1$ *thus* $A = -13/17$. *Consequently,* $B = 52/17 - 3 = (52 - 51)/17 = 1/17$. *Obviously* $C = 1/4$ *thus we find*

$$\boxed{y = c_1 + c_2 e^{-4x} - \frac{13}{17}\cos(x) + \frac{1}{17}\sin(x) + \frac{1}{4}x}.$$

We have enough examples to appreciate the theorem given below:

Theorem 3.4.14. *superposition principle for linear differential equations.*

Suppose $L[y] = 0$ is an n-th order linear differential equation with continuous coefficient functions on an interval I with fundamental solution set $S = \{y_1, y_2, \dots, y_n\}$ on I. Furthermore, suppose $L[y_{p_j}] = f_j$ for functions f_j on I then for any choice of constants $b_1, b_2, \dots, b_k$ the function $y = \sum_{j=1}^k b_j y_{p_j}$ forms the particular solution of $L[y] = \sum_{j=1}^k b_j f_j$ on the interval I.

Proof: we just use k-fold additivity and then homogeneity of L to show:

$$L\left[\sum_{j=1}^k b_j y_{p_j}\right] = \sum_{j=1}^k L[b_j y_{p_j}] = \sum_{j=1}^k b_j L[y_{p_j}] = \sum_{j=1}^k b_j f_j. \qquad \square$$

The Superposition Theorem paired with Theorem 3.2.23 allow us to find general solutions for complicated problems by breaking down the problem into pieces. In the example that follows we already dealt with the pieces in previous examples.

Example 3.4.15. **Problem:** *find the general solution of* $y'' + 4y' = 17(\cos(x) + 3\sin(x) + 1) + 16x = f$ *(introduced* f *for convenience here)*

Solution: *observe that* $L = D^2 + 4D$ *for both Example 3.4.12 and Example 3.4.13. We derived that the particular solutions* $y_{p_1} = \frac{1}{8}x^2 - \frac{1}{16}x$ *and* $y_{p_2} = -\frac{13}{17}\cos(x) + \frac{1}{17}\sin(x) + \frac{1}{4}x$ *satisfy*

$$L[y_{p_1}] = f_1 = x \qquad \& \qquad L[y_{p_2}] = f_2 = \cos(x) + 3\sin(x) + 1$$

[5] I know this by experience, but you can derive this by the annihilator method, of course the merit is made manifest in the successful selection of A, B below to actually solve $y_p'' + 4y_p' = x$.

Note that $f = 17f_2 + 16f_1$ *thus* $L[y] = f$ *has particular solution* $y = 17y_{p_2} + 16y_{p_1}$ *by the superposition principle. Therefore, the general solution is given by:*

$$y = c_1 + c_2e^{-4x} - 13\cos(x) + \sin(x) + \frac{17}{4}x + 2x^2 - x.$$

Or, collecting the x*-terms together,*

$$y = c_1 + c_2e^{-4x} - 13\cos(x) + \sin(x) + \frac{13}{4}x + 2x^2.$$

Example 3.4.16. Problem: *find the general solution of* $y'' + 5y' + 6y = 2\sinh(t)$

Solution: *It is easy to see that* $y'' + 5y' + 6y = e^t$ *has* $y_{p_1} = \frac{1}{12}e^t$. *On the other hand, it is easy to see that* $y'' + 5y' + 6y = e^{-t}$ *has solution* $y_{p_2} = \frac{1}{2}e^{-t}$. *The definition of hyperbolic sine gives* $2\sinh(t) = e^t - e^{-t}$ *hence, by the principle of superposition we find particular solution of* $y'' + 5y' + 6y = 2\sinh(t)$ *is simply* $y_p = y_{p_1} - y_{p_2}$. *Note* $\lambda^2 + 5\lambda + 6 = 0$ *factors to* $(\lambda + 2)(\lambda + 3) = 0$ *hence* $y_h = c_1e^{-2t} + c_2e^{-3t}$. *Therefore, the general solution of* $y'' + 5y' + 6y = 2\sinh(t)$ *is*

$$\boxed{y = c_1e^{-2t} + c_2e^{-3t} + \frac{1}{12}e^t - \frac{1}{2}e^{-t}}.$$

Naturally, you can solve the example above directly. I was merely illustrating the superposition principle.

Example 3.4.17. Problem: *find the general solution of* $y'' + 5y' + 6y = 2\sinh(t)$

Solution: *a natural choice for the particular solution is* $y_p = A\cosh(t) + B\sinh(t)$ *hence*

$$y_p' = A\sinh(t) + B\cosh(t), \quad y_p'' = A\cosh(t) + B\sinh(t) = y_p$$

Thus $y_p'' + 5y_p' + 6y_p = 5y_p' + 7y_p = 2\sinh(t)$ *and we find*

$$(7A + 5B)\cosh(t) + (7B + 5A)\sinh(t) = 2\sinh(t)$$

Thus $7A + 5B = 0$ *and* $7B + 5A = 2$. *Algebra yields* $A = -5/12$ *and* $B = 7/12$. *Therefore, as the characteristic values are* $\lambda = -2, -3$ *the general solution is given as follows:*

$$\boxed{y = c_1e^{-2x} + c_2e^{-3x} - \frac{5}{12}\cosh(t) + \frac{7}{12}\sinh(t)}.$$

I invite the reader to verify the answers in the previous pair of examples are in fact equivalent.

3.5 variation of parameters

The method of annihilators is deeply satisfying, but sadly most function escape its reach. For example, if the forcing function was $\sec(x)$ or $\tan(x)$ or $\ln(x)$ then we would be unable to annihilate these functions with some polynomial in D. Moreover, if the DEqn $L[y] = f$ has nonconstant coefficients then the problem of factoring L into linear factors $L_1, L_2, \dots, L_n$ is notoriously difficult[6]. If we had a factorization and a way to annihilate the forcing function we might be able to extend the method of the last section, but, this is not a particularly easy path to implement in any generality. In contrast, the technique of variation of parameters is both general and amazingly simple.

We begin by assuming the existence of a fundamental solution set for $L[y] = f$; assume $\{y_1, y_2, \dots, y_n\}$ is a linearly independent set of solutions for $L[y] = 0$. We **propose** the particular solution y_p can be written as a linear combination of the fundmental solutions with coefficients of functions $v_1, v_2, \dots, v_n$ (these are the "parameters")

$$y_p = v_1y_1 + v_2y_2 + \cdots + v_ny_n$$

Differentiate,

$$y_p' = v_1'y_1 + v_2'y_2 + \cdots + v_n'y_n + v_1y_1' + v_2y_2' + \cdots + v_ny_n'$$

Let constraint 1 state that $v_1'y_1+v_2'y_2+\cdots+v_n'y_n = 0$ and differentiate y_p' in view of this added constraint, once more we apply the product-rule n-fold times:

$$y_p'' = v_1'y_1' + v_2'y_2' + \cdots + v_n'y_n' + v_1y_1'' + v_2y_2'' + \cdots + v_ny_n''$$

Let constraint 2 state that $v_1'y_1' + v_2'y_2' + \cdots + v_n'y_n' = 0$ and differentiate y_p'' in view of contraints 1 and 2,

$$y_p''' = v_1'y_1'' + v_2'y_2'' + \cdots + v_n'y_n'' + v_1y_1''' + v_2y_2''' + \cdots + v_ny_n'''$$

Let constraint 3 state that $v_1'y_1'' + v_2'y_2'' + \cdots + v_n'y_n'' = 0$. We continue in this fashion adding constraints after each differentiation of the form $v_1'{y_1}^{(j)} + v_2'{y_2}^{(j)} + \cdots + v_n'{y_n}^{(j)} = 0$ for $j = 3, 4, \dots, n-2$. Note this brings us to

$$y_p^{(n-1)} = v_1y_1^{(n-1)} + v_2y_2^{(n-1)} + \cdots + v_ny_n^{(n-1)}.$$

Thus far we have given $(n-1)$-constraints on $[v_1', v_2', \dots, v_n']$. We need one more constraint to fix the solution. Remember we need $L[y_p] = f$; $a_ny_p^{(n)} + a_{n-1}y_p^{(n-1)} + \cdots + a_1y_p' + a_0y_p = f$ thus:

$$y_p^{(n)} = \frac{f}{a_n} - \frac{a_{n-1}}{a_n}y_p^{(n-1)} - \cdots - \frac{a_1}{a_n}y_p' - \frac{a_0}{a_n}y_p. \qquad (\star)$$

Differentiating $y_p = v_1y_1 + v_2y_2 + \cdots + v_ny_n$ and apply the previous contraints to obtain:

$$y_p^{(n)} = v_1'y_1^{(n-1)} + v_2'y_2^{(n-1)} + \cdots + v_n'y_n^{(n-1)} + v_1y_1^{(n)} + v_2y_2^{(n)} + \cdots + v_ny_n^{(n)}. \qquad (\star^2)$$

[6]we'll tackle the problem for the Cauchy Euler problem later this chapter, see Rabenstein for some more exotic examples of factorization of operators

Equate $\star$ and $\star^2$ to obtain:

$$\begin{aligned}
\frac{f}{a_n} &= \frac{a_{n-1}}{a_n} y_p^{(n-1)} + \cdots + \frac{a_1}{a_n} y_p' + \frac{a_0}{a_n} y_p + \\
&\quad + v_1' y_1^{(n-1)} + v_2' y_2^{(n-1)} + \cdots + v_n' y_n^{(n-1)} + v_1 y_1^{(n)} + v_2 y_2^{(n)} + \cdots + v_n y_n^{(n)} \\
&= v_1' y_1^{(n-1)} + v_2' y_2^{(n-1)} + \cdots + v_n' y_n^{(n-1)} \\
&\quad + \frac{a_n}{a_n}\left(v_1 y_1^{(n)} + v_2 y_2^{(n)} + \cdots + v_n y_n^{(n)} \right) + \\
&\quad + \frac{a_{n-1}}{a_n}\left(v_1 y_1^{(n-1)} + v_2 y_2^{(n-1)} + \cdots + v_n y_n^{(n-1)} \right) + \\
&\quad + \cdots + \\
&\quad + \frac{a_1}{a_n}\left(v_1 y_1' + v_2 y_2' + \cdots + v_n y_n' \right) + \\
&\quad + \frac{a_0}{a_n}\left(v_1 y_1 + v_2 y_2 + \cdots + v_n y_n \right) \\
&= v_1' y_1^{(n-1)} + v_2' y_2^{(n-1)} + \cdots + v_n' y_n^{(n-1)} \\
&\quad + \frac{v_1}{a_n}\left(a_n y_1^{(n)} + a_{n-1} y_1^{(n-1)} + \cdots + a_1 y_1' + a_0 y_1 \right) + \\
&\quad + \frac{v_2}{a_n}\left(a_n y_2^{(n)} + a_{n-1} y_2^{(n-1)} + \cdots + a_1 y_2' + a_0 y_2 \right) + \\
&\quad + \cdots + \\
&\quad + \frac{v_n}{a_n}\left(a_n y_n^{(n)} + a_{n-1} y_n^{(n-1)} + \cdots + a_1 y_n' + a_0 y_n \right) \\
&= v_1' y_1^{(n-1)} + v_2' y_2^{(n-1)} + \cdots + v_n' y_n^{(n-1)}.
\end{aligned}$$

In the step before the last we used the fact that $L[y_j] = 0$ for each y_j in the given fundamental solution set. With this calculation we obtain our n-th condition on the derivatives of the parameters. In total, we seek to impose

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix} \begin{bmatrix} v_1' \\ v_2' \\ \vdots \\ v_n' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ f/a_n \end{bmatrix}. \qquad (\star^3)$$

Observe that the coefficient matrix of the system above is the **Wronskian Matrix**. Since we assumed $\{y_1, y_2, \dots, y_n\}$ is a fundamental solution set we know that the Wronskian is nonzero which means the equation above has a unique solution. Therefore, the constraints we proposed are consistent and attainable for **any** n-th order linear ODE.

Let us pause to learn a little matrix theory convenient to our current endeavors. Nonsingular system of linear equations by Cramer's rule. To solve $A\vec{v} = \vec{b}$ you can follow the procedure below: to solve for v_k of $\vec{v} = (v_1, v_2, \dots, v_k, \dots, v_n)$ we

(1.) take the matrix A and replace the k-th column with the vector $\vec{b}$ call this matrix S_k

(2.) calculate $det(S_k)$ and $det(A)$

(3.) the solution is simply $v_k = \frac{det(S_k)}{det(A)}$.

Cramer's rule is a horrible method for specific numerical systems of linear equations[7]. But, it has for us the advantage of giving a nice, neat formula for the matrices of functions we consider here.

Example 3.5.1. *Suppose you want to solve $x+y+z=6$, $x+z=4$ and $y-z=-1$ simultaneously. Note in matrix notation we have:*

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ -1 \end{bmatrix}.$$

We can swap out columns $1,2$ and 3 to obtain S_1, S_2 and S_3

$$S_1 = \begin{bmatrix} 6 & 1 & 1 \\ 4 & 0 & 1 \\ -1 & 1 & -1 \end{bmatrix} \qquad S_2 = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 4 & 1 \\ 0 & -1 & -1 \end{bmatrix} \qquad S_3 = \begin{bmatrix} 1 & 1 & 6 \\ 1 & 0 & 4 \\ 0 & 1 & -1 \end{bmatrix}$$

You can calculate $det(S_1)=1$, $det(S_1)=2$ and $det(S_3)=3$. Likewise $det(A)=1$. Cramer's Rule states the solution is $x=\frac{det(S_1)}{det(A)}=1$, $y=\frac{det(S_2)}{det(A)}=2$ and $z=\frac{det(S_3)}{det(A)}=3$.

In the notation introduced above we see $\star^3$ has

$$A = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix} \qquad \& \qquad \vec{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ f/a_n \end{bmatrix}.$$

Once more define S_k as the matrix obtained by swapping the k-th column of A for the column vector $\vec{b}$ and let W be the Wronskian which is $det(A)$ in our current notation. We obtain the following solutions for $v_1', v_2', \dots, v_n'$ by Cramer's Rule:

$$v_1' = \frac{det(S_1)}{W},\ v_2' = \frac{det(S_2)}{W}, \dots,\ v_n' = \frac{det(S_n)}{W}$$

Finally, we can integrate to find the formulas for the parameters. Taking x as the independent parameter we note $v_k' = \frac{dv_k}{dx}$ hence:

$$v_1 = \int \frac{det(S_1)}{W}\,dx, \quad v_2 = \int \frac{det(S_2)}{W}\,dx, \ \dots, \ v_n = \int \frac{det(S_n)}{W}\,dx.$$

The matrix S_k has a rather special form and we can simplify the determinants above in terms of the so-called **sub-Wronskian** determinants. Define $W_k = W(y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_n; x)$ then it follows by Laplace's Expansion by minors formula that $det(S_k) = (-1)^{n+k}\frac{f}{a_n}W_k$. Thus,

$$\boxed{v_1 = \int (-1)^{n+1}\frac{fW_1}{a_nW}\,dx, \quad v_2 = \int (-1)^{n+2}\frac{fW_2}{a_nW}\,dx, \ \dots, \ v_n = \int \frac{fW_n}{a_nW}\,dx}.$$

Of course, you don't have to think about subWronskians, we could just use the formula in terms of $det(S_k)$. Include the subWronskain comment in part to connect with formulas given in Nagel Saff and Snider (Ritger & Rose does not have detailed plug-and-chug formulas on this problem, see page 154). In any event, we should now enjoy the spoils of this conquest. Let us examine how to calculate $y_p = v_1y_1 + \cdots + v_ny_n$ for particular n.

(1.) **(n=1)** $a_1\frac{dy}{dx} + a_0y = f$ has $W(y_1;x) = y_1$ and $W_1 = 1$. It follows that the solution $y = y_1v_1$ has $v_1 = \int \frac{f}{a_1y_1}\,dx$ where y_1 is the solution of $a_1\frac{dy}{dx} + a_0y = 0$ which is given by $y_1 = exp(\int \frac{-a_0}{a_1}\,dx)$. In other words, variation of parameters reduces to the integrating factor method[8] for $n=1$.

[7] Gaussian elimination is faster and more general, see my linear algebra notes or any text on the subject!

[8] note $\frac{dy}{dx} + \frac{a_0}{a_1}y = 0$ implies $I = exp(\int \frac{a_0}{a_1}\,dx)$ hence $\frac{d}{dx}(Iy) = 0$ and so $y = C/I$ and taking $C=1$ derives y_1.

(2.) (n=2) Suppose $a_2y'' + a_1y' + a_0y = f$ has fundamental solution set $\{y_1, y_2\}$ then

$$W = det \begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} = y_1y_2' - y_2y_1'$$

furthermore, calculate:

$$det(S_1) = det \begin{bmatrix} 0 & y_2 \\ f/a_2 & y_2' \end{bmatrix} = -\frac{fy_2}{a_2} \qquad \& \qquad det(S_2) = det \begin{bmatrix} y_1 & 0 \\ y_1' & f/a_2 \end{bmatrix} = \frac{fy_1}{a_2}$$

Therefore,

$$\boxed{v_1 = \int \frac{-fy_2}{a_2(y_1y_2' - y_2y_1')}\,dx \qquad \& \qquad v_2 = \int \frac{fy_1}{a_2(y_1y_2' - y_2y_1')}\,dx}$$

give the particular solution $y_p = v_1y_1 + v_2y_2$. Note that if the integrals above are indefinite then the general solution is given by:

$$\boxed{y = y_1 \int \frac{-fy_2}{a_2}\,dx + y_2 \int \frac{fy_1}{a_2}\,dx}.$$

Formulas for $n = 3, 4$ are tedious to derive and I leave them to the reader in the general case. Most applications involve $n = 2$.

Example 3.5.2. *Solve* $y'' + y = \sec(x)$. *The characteristic equation* $\lambda^2 + 1 = 0$ *yields* $\lambda = \pm i$ *hence* $y_1 = \cos(x), y_2 = \sin(x)$. *Observe the Wronskian simplifies nicely in this case:* $W = y_1y_2' - y_2y_1' = \cos^2(x) + \sin^2(x) = 1$. *Hence,*

$$v_1 = \int \frac{-fy_2}{W}\,dx = \int -\sec(x)\sin(x)\,dx = -\int \frac{\sin(x)}{\cos(x)}\,dx = \ln|\cos(x)| + c_1.$$

and,

$$v_2 = \int \frac{fy_1}{W}\,dx = \int \sec(x)\cos(x)\,dx = \int dx = x + c_2.$$

we find the general solution $y = y_1v_1 + y_2v_2$ *is simply:*

$$\boxed{y = c_1\cos(x) + c_2\sin(x) + \cos(x)\ln|\cos(x)| + x\sin(x)}.$$

Sometimes variation of parameters does not include the c_1, c_2 in the formulas for v_1 and v_2. In that case the particular solution truly is $y_p = y_1v_1 + y_2v_2$ and the general solution is found by $y = y_h + y_p$ where $y_h = c_1y_1 + c_2y_2$. Whatever system of notation you choose, please understand that in the end there must be a term $c_1y_1 + c_2y_2$ in the general solution.

Example 3.5.3. *Solve* $y'' - 2y' + y = f$. *The characteristic equation* $\lambda^2 - 2\lambda + 1 = (\lambda - 1)^2 = 0$ *yields* $\lambda_1 = \lambda_2 = 1$ *hence* $y_1 = e^x, y_2 = xe^x$. *Observe the Wronskian simplifies nicely in this case:* $W = y_1y_2' - y_2y_1' = e^x(e^x + xe^x) - e^xxe^x = e^{2x}$. *Hence,*

$$v_1 = \int \frac{-f(x)xe^x}{e^{2x}}\,dx \qquad \& \qquad v_2 = \int \frac{f(x)e^x}{e^{2x}}\,dx.$$

we find the general solution $y = y_1v_1 + y_2v_2$ *is simply:*

$$\boxed{y = c_1e^x + c_2xe^x - e^x \int \frac{f(x)xe^x}{e^{2x}}\,dx + xe^x \int \frac{f(x)e^x}{e^{2x}}\,dx}.$$

In particular, if $f(x) = e^x\sin(x)$ *then*

$$v_1 = \int -x\sin(x)\,dx = x\cos(x) - \sin(x) \qquad \& \qquad v_2 = \int \sin(x)\,dx = -\cos(x).$$

Hence, $y_p = (x\cos(x) - \sin(x))e^x + xe^x(-\cos(x)) = -e^x\sin(x)$. *The general solution is*

$$\boxed{y = c_1e^x + c_2xe^x - e^x\sin(x)}.$$

Notice that we could also solve $y'' - 2y' + y = e^x \sin(x)$ via the method of undetermined coefficients. In fact, any problem we can solve by undetermined coefficients we can also solve by variation of parameters. However, given the choice, it is usually easier to use undetermined coefficients.

Example 3.5.4. *Solve $y''' + y' = x\ln(x)$. The characteristic equation has $\lambda^3 + \lambda = \lambda(\lambda^2 + 1) = 0$ hence $\lambda = 0$ and $\lambda = \pm i$. The fundmental solutions are $y_1 = 1$, $y_2 = \cos(x)$, $y_3 = \sin(x)$. Calculate,*

$$W(1, \cos(x), \sin(x); x) = det \begin{bmatrix} 1 & \cos(x) & \sin(x) \\ 0 & -\sin(x) & \cos(x) \\ 0 & -\cos(x) & -\sin(x) \end{bmatrix} = 1(\sin^2(x) + \cos^2(x)) = 1.$$

Swapping the first column of the Wronskian matrix with $(0, 0, x\ln(x))$ gives us S_1 and we find

$$det(S_1) = det \begin{bmatrix} 0 & \cos(x) & \sin(x) \\ 0 & -\sin(x) & \cos(x) \\ x\ln(x) & -\cos(x) & -\sin(x) \end{bmatrix} = x\ln(x).$$

Swapping the second column of the Wronskian matrix with $(0, 0, x\ln(x))$ gives us S_2 and we find

$$det(S_2) = det \begin{bmatrix} 1 & 0 & \sin(x) \\ 0 & 0 & \cos(x) \\ 0 & x\ln(x) & -\sin(x) \end{bmatrix} = -x\ln(x)\cos(x).$$

Swapping the third column of the Wronskian matrix with $(0, 0, x\ln(x))$ gives us S_3 and we find

$$det(S_3) = \det \begin{bmatrix} 1 & \cos(x) & 0 \\ 0 & -\sin(x) & 0 \\ 0 & -\cos(x) & x\ln(x) \end{bmatrix} = -x\ln(x)\sin(x).$$

Note, integration by parts yields[9] *$v_1 = \int x\ln(x)\, dx = \frac{1}{2}x^2\ln(x) - \frac{1}{4}x^2$. The integrals of $v_2 = -\int x\ln(x)\cos(x)\, dx$ and $v_3 = -\int x\ln(x)\sin(x)\, dx$ are not elementary. However, we can express the general solution as:*

$$\boxed{y = c_1 + c_2\cos(x) + c_3\sin(x) + \frac{1}{2}x^2\ln(x) - \frac{1}{4}x^2 - \cos(x)\int x\ln(x)\cos(x)\, dx - \sin(x)\int x\ln(x)\sin(x)\, dx}.$$

If you use Mathematica directly, or Wolfram Alpha or other such software then some of the integrals will be given in terms of unusual functions such as *hypergeometric functions* or *polylogarithms* or *the cosine integral function* or the *exponential integral function* or the *sine integral function*, or *Bessel functions* and so forth... the list of nonstandard, but known, functions is very lenghthy at this point. What this means is that when you find an integral you cannot perform as part of an answer it may well be that the values of that integral are known, tabulated and often even automated as a built-in command. Moreover, if you randomly try other nonhomogeneous ODEs then you'll often find solutions appear in this larger class of named functions. More generally, the solutions appear as series of orthogonal functions. But, I suppose I am getting a little ahead of the story here. In the next section we explore substitutions of a particular sort for the n-th order problem.

[9] I am just calculating an antiderivative here since the homogeneous solution will account for the neccessary constants in the general solution

3.6 reduction of order

We return to the question of the homogeneous linear ODE $L[y] = 0$. Suppose we are **given** a solution y_1 with

$$a_n y_1^{(n)} + a_{n-1} y_1^{(n-1)} + \cdots + a_1 y_1' + a_0 y_1 = 0$$

on an interval I. To find a second solution we **propose** there exists v such that $y_2 = vy_1$ is a solution of $L[y] = 0$. I invite the reader to verify the following:

$$y_2' = v'y_1 + vy_1'$$

$$y_2'' = v''y_1 + 2v'y_1' + vy_1''$$

$$y_2''' = v'''y_1 + 3v''y_1' + 3v'y_1'' + vy_1'''$$

and, by an inductive argument, we arrive at

$$y_2^{(n)} = v^{(n)} y_1 + nv^{(n-1)} y_1' + \cdots + nv' y_1^{(n-1)} + vy_1^{(n)}$$

where the other coefficients are the binomial coefficients. I suppose it's worth mentioning the formula below is known as *Leibniz' product formula:*

$$\boxed{\frac{d^n}{dx^b}\Big[F(x)\,G(x)\Big] = \sum_{k=0}^{n} \binom{n}{k} F^{(n-k)}(x) G^{(k)}(x)}$$

Returning to the substitution $y_2 = vy_1$ we find that the condition $L[y_2] = 0$ gives

$$a_n(vy_1)^{(n)} + \cdots + a_1(vy_1)' + a_0 vy_1 = 0$$

Thus,

$$a_n\big[v^{(n)} y_1 + nv^{(n-1)} y_1' + \cdots + nv' y_1^{(n-1)} + vy_1^{(n)}\big] + \cdots + a_1\big[v'y_1 + vy_1'\big] + a_0 vy_1 = 0$$

Notice how all the terms with v collect together to give $v\big[y_1^{(n)} + \cdots + a_{n-1}y_1' + a_n y_1\big]$ which vanishes since y_1 is a solution. Therefore, the equation $L[y_2] = 0$ reduces to:

$$a_n\big[v^{(n)} y_1 + nv^{(n-1)} y_1' + \cdots + nv' y_1^{(n-1)}\big] + \cdots + a_1 v' y_1 = 0$$

If we substitute $z = v'$ then the equation is clearly an $(n-1)$-th order linear ODE for z;

$$a_n\big[z^{(n-1)} y_1 + nz^{(n-2)} y_1' + \cdots + nzy_1^{(n-1)}\big] + \cdots + a_1 zy_1 = 0.$$

I include this derivation to show you that the method extends to the n-th order problem. However, we are primarily interested in the $n = 2$ case. In that particular case we can derive a nice formula for y_2.

3.6.1 the second linearly independent solution formula

Let a, b, c be functions and suppose $ay'' + by' + cy = 0$ has solution y_1. Let $y_2 = vy_1$ and seek a formula for v for which y_2 is a solution of the $ay'' + by' + cy = 0$. Substitute $y_2 = vy_1$ and differentiate the product,

$$a(v''y_1 + 2v'y_1' + vy_1'') + b(v'y_1 + vy_1') + cvy_1 = 0$$

Apply $ay_1'' + by_1' + cy_1 = 0$ to obtain:

$$a(v''y_1 + 2v'y_1') + bv'y_1 = 0$$

Now let $z = v'$ thus $z' = v''$

$$ay_1z' + 2ay_1'z + by_1z = 0 \quad \Rightarrow \quad \frac{dz}{dx} + \left[\frac{2ay_1' + by_1}{ay_1}\right]z = 0.$$

Apply the integrating factor method with $I = exp(\int \frac{2ay_1'+by_1}{ay_1} dx)$ we find

$$\frac{d}{dx}\left[Iz\right] = 0 \quad \Rightarrow \quad Iz = C \quad \Rightarrow \quad z = \frac{C}{I} = Cexp\left(-\int \frac{2ay_1' + by_1}{ay_1} dx\right)$$

Recall $z = \frac{dv}{dx}$ thus we integrate to find $v = \int Cexp\left(-\int \frac{2ay_1'+by_1}{ay_1} dx\right) dx$ thus

$$y_2 = y_1 \int Cexp\left(-\int \frac{2ay_1' + by_1}{ay_1} dx\right) dx$$

It is convenient to take $C = 1$ since we are just seeking a particular function to construct the solution set. Moreover, notice that the integral $\int \frac{-2}{y_1}\frac{dy_1}{dx}\, dx = -2\ln|y_1| = \ln(1/y_1^2)$ thus it follows

$$\boxed{y_2 = y_1 \int \frac{1}{y_1^2} exp\left(-\int \frac{b}{a}\, dx\right) dx} \tag{3.2}$$

Example 3.6.1. *Consider $y'' - 2y' + y = 0$. We found $y_1 = e^x$ by making a simple guess of $y = e^{\lambda x}$ and working out the algebra. Let us now find how to derive y_2 in view of the derivation preceding this example. Identify $a = 1, b = -2, c = 1$. Suppose $y_2 = vy_1$. We found that*

$$y_2 = y_1 \int \frac{1}{y_1^2} exp\left(-\int \frac{b}{a}\, dx\right) dx = e^x \int \frac{1}{e^{2x}} exp\left(\int 2\, dx\right) dx = e^x \int \frac{e^{2x}}{e^{2x}}\, dx = e^x \int dx$$

Thus $y_2 = xe^x$.

This example should suffice for the moment. We will use this formula in a couple other places. Notice if we have some method to find at least one solution for $ay'' + by' + cy = 0$ then this formula allows us to find a second, linearly independent[10] solution.

[10] no, I have not proved this, perhaps you could try

3.7 operator factorizations

In this section we consider a method to solve $L[y] = f$ given that $L = L_1L_2 \cdots L_n$ and L_j are all first order differential operators. Without loss of generality this means $L_j = a_jD + b_j$ for $j = 1, 2, \ldots, n$. We do not suppose these operators commute. Let $z_1 = (L_2L_3 \cdots L_n)[y]$ and note that in z_1 the n-th order ODE for y simplifies to

$$L_1[z_1] = f \quad \Rightarrow \quad \frac{dz_1}{dx} + \frac{b_1}{a_1}z_1 = f \quad \Rightarrow \quad exp\left[\int \frac{b_1}{a_1}dx\right]\frac{dz_1}{dx} + exp\left[\int \frac{b_1}{a_1}dx\right]\frac{b_1}{a_1}z_1 = \frac{f}{a_1}exp\left[\int \frac{b_1}{a_1}dx\right]$$

Consequently,

$$\frac{d}{dx}\left[z_1\, exp\left[\int \frac{b_1}{a_1}dx\right]\right] = \frac{f}{a_1}exp\left[\int \frac{b_1}{a_1}dx\right]$$

integrating and solving for z_1 yields:

$$z_1 = exp\left[-\int \frac{b_1}{a_1}dx\right]\left[c_1 + \int \frac{f}{a_1}exp\left[\int \frac{b_1}{a_1}dx\right]dx\right]$$

Next let $z_2 = (L_3 \cdots L_n)[y]$ observe that $L_1(L_2[z_2]) = f$ implies $z_1 = L_2[z_2]$ hence we should solve

$$a_2\frac{dz}{dx} + b_2z_2 = z_1$$

By the calculation for z_1 we find, letting z_1 play the role f did in the previous calculation,

$$z_2 = exp\left[-\int \frac{b_2}{a_2}dx\right]\left[c_2 + \int \frac{z_1}{a_2}exp\left[\int \frac{b_2}{a_2}dx\right]dx\right]$$

Well, I guess you see where this is going, let $z_3 = (L_4 \cdots L_n)[y]$ and observe $(L_1L_2)[L_3[z_3]] = f$ hence $L_3[z_3] = z_2$. We must solve $a_3z_3' + b_3z_3 = z_2$ hence

$$z_3 = exp\left[-\int \frac{b_3}{a_3}dx\right]\left[c_3 + \int \frac{z_2}{a_3}exp\left[\int \frac{b_3}{a_3}dx\right]dx\right].$$

Eventually we reach $y = z_n$ where $(L_1L_2 \cdots L_n)[z_n] = f$ and $a_nz_n' + b_nz_n = z_{n-1}$ will yield

$$y = exp\left[-\int \frac{b_n}{a_n}dx\right]\left[c_n + \int \frac{z_{n-1}}{a_n}exp\left[\int \frac{b_n}{a_n}dx\right]dx\right].$$

If we expand $z_{n-1}, z_{n-2}, \ldots, z_2, z_1$ we find the formula for the general solution of $L[y] = f$.

The trouble with this method is that its starting point is a factored differential operator. Many problems do not enjoy this structure from the outset. We have to do some nontrivial work to massage an arbitrary problem into this factored form. Rabenstein[11] claims that it is always possible to write L in factored form, but even in the $n = 2$ case the problem of factoring L is as difficult, if not more, then solving the differential equation!

[11] page 70, ok technically he only claims $n = 2$, I haven't found a general reference at this time

Example 3.7.1. *Let $L_1 = x\frac{d}{dx}$ and suppose $L_2 = 1 + \frac{d}{dx}$. Solve $(L_1L_2)[y] = 3$. We want to solve*

$$x\frac{d}{dx}\left[y + \frac{dy}{dx}\right] = 3$$

Let $z = y + \frac{dy}{dx}$ and consider

$$x\frac{dz}{dx} = 3 \quad \Rightarrow \quad \int dz = \int \frac{3dx}{x} \quad \Rightarrow \quad z = 3\ln|x| + c_1.$$

Hence solve,

$$y + \frac{dy}{dx} = 3\ln|x| + c_1$$

Multiply by integrating factor e^x and after a short calculation we find

$$y = e^{-x}\int \left[3\ln|x|e^x + c_1e^x\right] dx$$

Therefore,

$$y = c_2e^{-x} + c_1 + e^{-x}\int \left[3\ln|x|e^x\right] dx$$

Identify the fundamental solution set of $y_1 = e^{-x}$ and $y_2 = 1$. Note that $L_2[e^{-x}] = 0$ and $L_1[1] = 0$.

Curious, we just saw a non-constant coefficient differential equation which has the same fundamental solution set as $y'' + y' = 0$. I am curious how the solution will differ if we reverse the order of L_1 and L_2

Example 3.7.2. *Let L_1, L_2 be as before and solve $(L_2L_1)[y] = 3$. We want to solve*

$$\left[1 + \frac{d}{dx}\right]\left[x\frac{dy}{dx}\right] = 3$$

Let $z = x\frac{dy}{dx}$ and seek to solve $z + \frac{dz}{dx} = 3$. This is a constant coefficient ODE with $\lambda = -1$ and it is easy to see that $z = 3 + c_1e^{-x}$. Thus consider, $x\frac{dy}{dx} = 3 + c_1e^{-x}$ yields $dy = \left(\frac{3}{x} + c_1\frac{e^{-x}}{x}\right)dx$ and integration yields:

$$y = c_2 + c_1\int \frac{e^{-x}}{x}\,dx + 3\ln|x|.$$

The fundamental solution set has $y_1 = 1$ and $y_2 = \int \frac{e^{-x}}{x}\,dx$.

You can calculate that $L_1L_2 \neq L_2L_1$. This is part of what makes the last pair of examples interesting. On the other hand, perhaps you can start to appreciate the constant coefficient problem. In the next section we consider the next best thing; the *equidimensional* or *Cauchy Euler* problem. It turns out we can factor the differential operator for a Cauchy -Euler problem into commuting differential operators. This makes the structure of the solution set easy to catalogue.

3.8 cauchy euler problems

The general Cauchy-Euler problem is specified by n-constants $a_1, a_2, \dots, a_n$. If L is given by

$$L = x^n D^n + a_1 x^{n-1} D^{n-1} + \cdots + a_{n-1} x D + a_n$$

then $L[y] = 0$ is a Cauchy-Euler problem. Suppose the solution is of the form $y = x^R$ for some constant R. Note that

$$xD[x^R] = xRx^{R-1} = Rx^R$$
$$x^2 D^2[x^R] = x^2 R(R-1)x^{R-2} = R(R-1)x^R$$
$$x^3 D^3[x^R] = x^3 R(R-1)(R-2)x^{R-3} = R(R-1)(R-2)x^R$$
$$x^n D^n[x^R] = x^n R(R-1)(R-2)\cdots(R-n)x^{R-n} = R(R-1)(R-2)\cdots(R-n)x^R$$

Substitute into $L[y] = 0$ and obtain:

$$\Big(R(R-1)(R-2)\cdots(R-n) + a_1 R(R-1)(R-2)\cdots(R-n+1) + \cdots + a_{n-1}R + a_n\Big)x^R = 0$$

It follows that R must satisfy the **characteristic equation**

$$\boxed{R(R-1)(R-2)\cdots(R-n) + a_1 R(R-1)(R-2)\cdots(R-n+1) + \cdots + a_{n-1}R + a_n = 0.}$$

Notice that it is not simply obtained by placing powers of R next to the coefficients $a_1, a_2, \dots, a_n$. However, we do obtain an n-th order polynomial equation for R and it follows that we generally have n-solutions, some repeated, some complex. This begs an interesting question: *what does x to a complex power mean ?*

Definition 3.8.1. *complex power function with real base.*

Let $a, b \in \mathbb{R}$, define $x^{a+ib} = x^a(\cos(b\ln(x)) + i\sin(b\ln(x)))$.

Motivation: $x^c = e^{log(x^c)} = e^{clog(x)} = e^{a\ln(x)+ib\ln(x)} = e^{a\ln(x)}e^{ib\ln(x)} = x^a e^{ib\ln(x)}$.

I invite the reader to check that the power-rule holds for complex exponents:

Proposition 3.8.2. *let $c \in \mathbb{C}$ then for $x > 0$,*

$$\frac{d}{dx}\big[x^c\big] = cx^{c-1}.$$

This make a nice homework problem. I worked the analogous problem for the complex exponential in an earlier section. The rules for complexified problems apply here, the real and imaginary parts of the complex solution give us a pair of real solutions.

Example 3.8.3. *Solve $x^2y'' + xy' + y = 0$. Let $y = x^R$ then we must have*

$$R(R-1) + R + 1 = 0 \quad \Rightarrow \quad R^2 + 1 = 0 \quad \Rightarrow \quad R = \pm i$$

Hence $y = x^i$ is a complex solution. We defined, for $x > 0$, the complex power function $x^c = e^{c\ln(x)}$ hence

$$x^i = e^{i\ln(x)} = \cos(\ln(x)) + i\sin(\ln(x))$$

The real and imaginary parts of x^i give real solutions for $x^2y'' + xy' + y = 0$. We find

$$\boxed{y = c_1\cos(\ln(x)) + c_2\sin(\ln(x))}$$

Example 3.8.4. *Solve $x^2y'' + 4xy' + 2y = 0$. Let $y = x^R$ then we must have*

$$R(R-1) + 4R + 2 = 0 \quad \Rightarrow \quad R^2 + 3R + 2 = (R+1)(R+2) \quad \Rightarrow \quad R = -1, -2.$$

We find fundamental solutions $y_1 = 1/x$ and $y_2 = 1/x^2$ hence the general solution is

$$\boxed{y = c_1 \frac{1}{x} + c_2 \frac{1}{x^2}}$$

Example 3.8.5. *Solve $x^2y'' - 3xy' + 4y = 0$ for $x > 0$. Let $y = x^R$ then we must have*

$$R(R-1) - 3R + 4 = 0 \quad \Rightarrow \quad R^2 - 4R + 4 = (R-2)^2 \quad \Rightarrow \quad R = 2, 2.$$

We find fundamental solution $y_1 = x^2$. To find y_2 we must use another method. We derived that the second solution of $ay'' + by' + cy = 0$ can be found from the first via Equation 3.2:

$$y_2 = y_1 \int \frac{1}{y_1^2} exp\left(\int \frac{-b}{a} dx\right) dx.$$

In this problem identify that $a = x^2$ and $b = -3x$ whereas $y_1 = x^2$ and $y_1' = 2x$ thus:

$$y_2 = x^2 \int \frac{1}{x^4} exp\left(\int \frac{3x}{x^2} dx\right) dx = x^2 \int \frac{1}{x^4} exp\Big(3\ln(x)\Big) dx = x^2 \int \frac{dx}{x} = x^2 \ln(x).$$

The general solution is $\boxed{y = c_1x^2 + c_2x^2\ln(x)}$.

The first order case is also interesting:

Example 3.8.6. *Solve $x\frac{dy}{dx} - ay = 0$. Let $y = x^R$ and find $R - a = 0$ hence* $\boxed{y = c_1x^a}$. *The operator $xD - a$ has characteristic equation $R - a = 0$ hence the characteristic value is $R = a$.*

Let us take two first order problems and construct a second order problem. Notice the operator in the last example is given by $xD - a$. We compose two such operators to construct,

$$(xD - a)(xD - b)[y] = 0$$

We can calculate,

$$(xD - a)[xy' - by] = xD[xy' - by] - axy' + aby = xy' + x^2y'' - bxy' - axy' + aby$$

In operator notation we find

$$(xD - a)(xD - b) = x^2D^2 + (1 - a - b)xD + ab$$

from which it is clear that $(xD - a)(xD - b) = (xD - b)(xD - a)$. Moreover,

$$(xD - a)(xD - b)[y] = 0 \qquad \Leftrightarrow \qquad (x^2D^2 + (1 - a - b)xD + ab)[y] = 0$$

Example 3.8.7. *To construct an cauchy-euler equation with characteristic values of $a = 2+3i$ and $b = 2-3i$ we simply note that $1 - a - b = -3$ and $ab = 4 + 9 = 13$. We can check that the cauchy-euler problem $x^2y'' - 3xy' + 13y = 0$ has complex solutions $y = x^{2\pm 3i}$, suppose $y = x^R$ then it follows that R must solve the characteristic equation:*

$$R(R-1) - 3R + 13 = R^2 - 4R + 13 = (R-2)^2 + 9 = 0 \quad \Rightarrow \quad R = 2 \pm 3i.$$

Note $x^{2+3i} = e^{(2+3i)\ln(x)} = e^{ln(x^2)}e^{3i\ln(x)} = x^2(\cos(3\ln(x)) + i\sin(3\ln(x)))$ (you can just memorize it as I defined it, but these steps help me remember how this works) Thus, the DEqn $x^2y'' - 3xy' + 13y = 0$ has general solution

$$\boxed{y = c_1x^2\cos(3\ln(x)) + c_2x^2\sin(3\ln(x))}$$

We saw that noncommuting operators are tricky to work with in a previous section. Define $[L_1, L_2] = L_1L_2 - L_2L_1$ and note that $L_2L_1 = L_1L_2 - [L_1, L_2]$. The $[L_1, L_2]$ is called the **commutator**, when it is zero then the inputs to $[,]$ are said to commute. If you think about the homogeneous problem $(L_1L_2)[y] = 0$ then contrast with $(L_2L_1)[y] = 0$ we can understand why these are not the same in terms of the commutator. For example, suppose $L_2[y] = 0$ then it is clearly a solution of $(L_1L_2)[y] = 0$ since $L_1[L_2[y]] = L_1[0] = 0$. On the other hand,

$$(L_2L_1)[y] = (L_1L_2 - [L_1, L_2])[y] = L_1[L_2[y]] - [L_1, L_2][y] = -[L_1, L_2][y]$$

and there is no reason in general for the solution to vanish on the commutator above. If we could factor a given differential operator into **commuting** operators $L_1, L_2, \dots, L_n$ then the problem $L[y] = 0$ nicely splits into n-separate problems $L_1[y] = 0, L_2[y] = 0, \dots, L_n[y] = 0$.

With these comments in mind return to the question of solving $L[y] = 0$ for

$$L = x^nD^n + a_1x^{n-1}D^{n-1} + \cdots + a_{n-1}xD + a_n$$

note in the case $n = 2$ we can solve $R(R-1) + a_R + a_o = 0$ for solutions a, b and it follows that

$$x^2D^2 + a_1xD + a_2 = (xD - a)(xD - b)[y] = 0$$

The algebra to state a, b as functions of a_1, a_2 is a quadratic equation. Notice that for the third order operator it starts to get ugly, the fourth unpleasant, and the fifth, impossible in closed form for an arbitrary equidimensional quintic operator.

All of this said, I think it is at least possible to *explicitly*[12] factor the operator whenever we can factor the characteristic equation. Suppose R_1 is a solution to the characteristic equation hence $y_1 = x^{R_1}$ is a solution of $L[y] = 0$. I claim you can argue that $L_1 = (xD - a_1)$ is a factor of L. Likewise, for the other zeros $a_2, a_3, \dots, a_n$ the linear differential operators $L_j = (xD - a_j)$ must somehow appear as a factor of L. Hence we have n-first order differential operators and since I wrote $L = x^nD^n + \cdots + a_n$ it follows that $L = L_1L_2\cdots L_n$. From a DEqns perspective this discussion is not terribly useful as the process of factoring L into a polynomial in xD is not so intutive. Even the $n = 2$ case is tricky:

$$\begin{aligned}(xD - a)(xD - b)[f] &= (x^2D^2 + (1 - a - b)xD + ab)[f] \\ &= (xD - a)(xD - b)[f] \\ &= xD(xD[f] - b[f]) - axD[f] + ab[f] \\ &= \big((xD)^2 - (a + b)(xD) + ab\big)\,[f]\end{aligned}$$

Notice the polynomials in xD behave nicely but the x^2D^2 term does not translate simply into the xD formulas. Let's see if we can derive some general formula to transform x^nD^n into some polynomial in xD.

Calculate, for f a suitably differentiable function,

$$(xD)^2[f] = xD[xD[f]] = xD[xf'] = xf' + x^2f'' = (xD + x^2D^2)[f] \quad \Rightarrow \quad \boxed{x^2D^2 = (xD)^2 - xD}$$

Next, order three, using Leibniz' product rule for second derivative of a product,

$$\begin{aligned}(xD)^3[f] = (xD + x^2D^2)[xf'] &= xf' + x^2f'' + x^2(x''f' + 2x'f'' + xf''') \\ &= (xD + x^2D^2 + 2x^2D^2 + x^3D^3)[f] \\ &= (xD + 3x^2D^2 + x^3D^3)[f] \\ &= (xD + 3(xD)^2 - 3xD + x^3D^3)[f] \\ &= (3(xD)^2 - 2xD + x^3D^3)[f] \\ &\Rightarrow \boxed{x^3D^3 = (xD)^3 - 3(xD)^2 + 2xD}.\end{aligned}$$

[12]theoretically it is always possible by the fundamental theorem of algebra applied to the characteristic equation and the scheme I am about to outline

It should be fairly clear how to continue this to higher orders. Let's see how this might be useful[13] in the context of a particular third order cauchy-euler problem.

Example 3.8.8. *Solve* $(x^3D^3 + 3x^2D^2 + 2xD)[y] = 0$. *I'll use operator massage. By the calculations preceding this example:*

$$x^3D^3 + 3x^2D^2 + 2xD = (xD)^3 - 3(xD)^2 + 2xD + 3(xD)^2 - 3xD + 2xD = (xD)^3 + (xD)$$

Now I can do algebra since xD *commutes with itself,*

$$(xD)^3 + (xD) = xD((xD)^2 + 1) = xD(xD - i)(xD + i)$$

Hence $R = 0, R = \pm i$ *are evidentally the characteristic values and we find real solution*

$$\boxed{y = c_1 + c_2\cos(\ln(x)) + c_3\sin(\ln(x))}$$

Let's check this operator-based calculation against our characteristic equation method:

$$R(R-1)(R-2) + 3R(R-1) + 2R = R^3 - 3R^2 + 2R + 3R^2 - 3R + 2R = R^3 + R.$$

Which would then lead us to the same solution as we uncovered from the xD *factorization.*

If you're enjoying this section then you might want to read the paper *A Generalized Method of Undetermined Coefficients* which I wrote with my brother Dr. William Cook for Volume 15 of the CODEE Journal in 2022. See (click here for the CODEE journal paper). We generalize the annhilator method to solve a wide variety of differential equations. The genesis of that paper was basically a student asking my brother if it was possible to solve non-homogenous Cauchy Euler problems in the same fashion as the constant coefficient problems. In short, the answer is yes.

3.9 applications

We explore two interesting applications in this section:

1. springs with friction
2. RLC circuits

We begin by studying the homogeneous case and then add external forces (1.) or a voltage source (2.). The mathematics is nearly the same for both applications. Finally we study resonance.

3.9.1 springs with and without damping

Suppose a mass m undergoes one-dimensional motion under the influence of a spring force $F_s = -kx$ and a velocity dependent friction force $F_f = -\beta\dot{x}$. Newton's Second Law states $m\ddot{x} = -kx - \beta\dot{x}$. We find

$$\boxed{m\ddot{x} + \beta\dot{x} + kx = 0}$$

The constants m, β, k are non-negative and we assume $m \neq 0$ in all cases. Technically the value of m should be assigned kg, that of β should be assigned kg/s and the spring constant k should have a value with units of the form N/m. Please understand these are omitted in this section. When faced with a particular problem make sure you use quantities which have compatible units.

[13] my point in these calculations is not to find an optimal method to solve the cauchy euler problem, probably the characteristic equation is best, my point here is to explore the structure of operators and test our ability to differentiate and think!

Example 3.9.1. Problem: the over-damped spring: *Suppose $m = 1, \beta = 3$ and $k = 2$. If the mass has velocity $v = -2$ and position $x = 1$ when $t = 0$ then what is the resulting equaiton of motion?*

Solution: *We are faced with $\ddot{x} + 3\dot{x} + 2x = 0$. This gives chararcteristic equation $\lambda^2 + 3\lambda + 2 = 0$ hence $(\lambda+1)(\lambda+2) = 0$ thus $\lambda_1 = -1$ and $\lambda_2 = -2$ and the general solution is*

$$x(t) = c_1 e^{-t} + c_2 e^{-2t}$$

Note that $\dot{x}(t) = -c_1 e^{-t} - 2c_2 e^{-2t}$. Apply the given initial conditions,

$$x(0) = c_1 + c_2 = 1 \qquad \& \qquad \dot{x}(0) = -c_1 - 2c_2 = -2$$

You can solve these equations to obtain $c_2 = 1$ and $c_1 = 0$. Therefore, $\boxed{x(t) = e^{-2t}}$.

Example 3.9.2. Problem: the critically-damped spring: *Suppose $m = 1, \beta = 4$ and $k = 4$. If the mass has velocity $v = 1$ and position $x = 3$ when $t = 0$ then what is the resulting equaiton of motion?*

Solution: *We are faced with $\ddot{x} + 4\dot{x} + 4x = 0$. This gives chararcteristic equation $\lambda^2 + 4\lambda + 4 = 0$ hence $(\lambda+2)^2 = 0$ thus $\lambda_1 = \lambda_2 = -2$ and the general solution is*

$$x(t) = c_1 e^{-2t} + c_2 t e^{-2t}$$

Note that $\dot{x}(t) = -2c_1 e^{-t} + c_2(e^{-2t} - 2te^{-2t})$. Apply the given initial conditions,

$$x(0) = c_1 = 3 \qquad \& \qquad \dot{x}(0) = -2c_1 + c_2 = 1$$

You can solve these equations to obtain $c_1 = 3$ and $c_2 = 7$. Therefore, $\boxed{x(t) = 3e^{-2t} + 7te^{-2t}}$.

Example 3.9.3. Problem: the under-damped spring: *Suppose $m = 1, \beta = 2$ and $k = 6$. If the mass has velocity $v = 1$ and position $x = 1$ when $t = 0$ then what is the resulting equaiton of motion?*

Solution: *We are faced with $\ddot{x} + 2\dot{x} + 6x = 0$. This gives chararcteristic equation $\lambda^2 + 2\lambda + 6 = 0$ hence $(\lambda+1)^2 + 5 = 0$ thus $\lambda = -1 \pm i\sqrt{5}$ and the general solution is*

$$x(t) = c_1 e^{-t}\cos(\sqrt{5}\,t) + c_2 e^{-t}\sin(\sqrt{5}\,t)$$

Note that $\dot{x}(t) = c_1 e^{-t}(-\cos(\sqrt{5}\,t) - \sqrt{5}\sin(\sqrt{5}\,t) + c_2 e^{-t}(-\sin(\sqrt{5}\,t) + \sqrt{5}\cos(\sqrt{5}\,t)$. Apply the given initial conditions,

$$x(0) = c_1 = 1 \qquad \& \qquad \dot{x}(0) = -c_1 + \sqrt{5}c_2 = 1$$

You can solve these equations to obtain $c_1 = 1$ and $c_2 = 2/\sqrt{5}$. Therefore,

$$\boxed{x(t) = e^{-t}\cos(\sqrt{5}\,t) + \frac{2}{\sqrt{5}}e^{-t}\sin(\sqrt{5}\,t)}.$$

Example 3.9.4. Problem: spring without damping; simple harmonic oscillator: *Suppose $\beta = 0$ and m, k are nonzero. If the mass has velocity $v(0) = v_o$ and position $x(0) = x_o$ then find the resulting equaiton of motion.*

Solution: *We are faced with $m\ddot{x} + kx = 0$. This gives chararcteristic equation $m\lambda^2 + k = 0$ hence $\lambda = \pm i\sqrt{\frac{k}{m}}$ and the general solution is, using $\omega = \frac{k}{m}$,*

$$x(t) = c_1\cos(\omega t) + c_2\sin(\omega\, t)$$

Note that

$$\dot{x}(t) = -c_1\omega\sin(\omega\, t) + c_2\omega\cos(\omega\, t).$$

Apply the given initial conditions,

$$x(0) = c_1 = x_o \qquad \& \qquad \dot{x}(0) = c_2\omega = v_o$$

Therefore,

$$\boxed{x(t) = x_o\cos(\omega t) + \frac{v_o}{\omega}\sin(\omega\, t)}.$$

3.9.2 the RLC-circuit

Now we turn to circuits. Suppose a resistor R, an inductor L and a capacitor C are placed in series then the we know that $V_R = IR$ by Ohm's Law for the resistor, whereas the voltage dropped on an inductor is proportional to the change in the current according to the definition of inductance paired with Faraday's Law: $V_L = L\frac{dI}{dt}$ for the inductor, the capacitor C has charge $\pm Q$ on its plates when $V_C = Q/C$. We also know $I = \frac{dQ}{dt}$ since the capacitor is in series with R and L. Finally, we apply Kirchoff's voltage law around the circuit to obtain $V_R + V_L + V_C = 0$, this yields:

$$IR + L\frac{dI}{dt} + \frac{Q}{C} = 0 \quad \Rightarrow \quad \boxed{L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = 0}.$$

Obviously there is an analogy to be made here:

$$\boxed{m \leftrightarrows L \qquad \beta \leftrightarrows R \qquad k \leftrightarrows \frac{1}{C}}$$

I will exploit this analogy to construct the following examples.

Example 3.9.5. Problem: the over-damped RLC circuit: *Suppose $L = 1, R = 3$ and $C = 1/2$. If the circuit has current $I = -2$ and charge $Q = 1$ when $t = 0$ then what is the charge as a function of time? What is the current at a function of time?*

Solution: *We are faced with $\ddot{Q} + 3\dot{Q} + 2Q = 0$. This gives chararcteristic equation $\lambda^2 + 3\lambda + 2 = 0$ hence $(\lambda+1)(\lambda+2) = 0$ thus $\lambda_1 = -1$ and $\lambda_2 = -2$ and the general solution is*

$$Q(t) = c_1e^{-t} + c_2e^{-2t}$$

Note that $\dot{Q}(t) = -c_1e^{-t} - 2c_2e^{-2t}$. Apply the given initial conditions,

$$Q(0) = c_1 + c_2 = 1 \qquad \& \qquad \dot{Q}(0) = -c_1 - 2c_2 = -2$$

You can solve these equations to obtain $c_2 = 1$ and $c_1 = 0$. Therefore, $\boxed{Q(t) = e^{-2t}}$. Differentiate to obtain the current $\boxed{I(t) = -2e^{-2t}}$.

Example 3.9.6. Problem: the critically-damped RLC circuit: *Suppose $L = 1, R = 4$ and $C = 1/4$. If the circuit has current $I = 1$ and charge $Q = 3$ when $t = 0$ then what is the charge as a function of time? What is the current at a function of time?*

Solution: *We are faced with $\ddot{Q} + 4\dot{Q} + 4Q = 0$. This gives chararcteristic equation $\lambda^2 + 4\lambda + 4 = 0$ hence $(\lambda+2)^2 = 0$ thus $\lambda_1 = \lambda_2 = -2$ and the general solution is*

$$Q(t) = c_1e^{-2t} + c_2te^{-2t}$$

Note that $\dot{Q}(t) = -2c_1e^{-t} + c_2(e^{-2t} - 2te^{-2t})$. Apply the given initial conditions,

$$Q(0) = c_1 = 3 \qquad \& \qquad \dot{Q}(0) = -2c_1 + c_2 = 1$$

You can solve these equations to obtain $c_1 = 3$ and $c_2 = 7$. Therefore, $\boxed{Q(t) = 3e^{-2t} + 7te^{-2t}}$. Differentiate the charge to find the current $\boxed{I(t) = e^{-2t} - 14te^{-2t}}$.

Example 3.9.7. Problem: the under-damped RLC circuit: *Suppose $L = 1, R = 2$ and $C = 1/6$. If the circuit has current $I = 1$ and charge $Q = 1$ when $t = 0$ then what is the charge as a function of time? What is the current at a function of time?*

Solution: *We are faced with $\ddot{Q} + 2\dot{Q} + 6Q = 0$. This gives chararcteristic equation $\lambda^2 + 2\lambda + 6 = 0$ hence $(\lambda + 1)^2 + 5 = 0$ thus $\lambda = -1 \pm i\sqrt{5}$ and the general solution is*

$$Q(t) = c_1 e^{-t}\cos(\sqrt{5}\,t) + c_2 e^{-t}\sin(\sqrt{5}\,t)$$

Note that $\dot{Q}(t) = c_1 e^{-t}\left(-\cos(\sqrt{5}\,t) - \sqrt{5}\sin(\sqrt{5}\,t)\right) + c_2 e^{-t}\left(-\sin(\sqrt{5}\,t) + \sqrt{5}\cos(\sqrt{5}\,t)\right)$. Apply the given initial conditions,

$$Q(0) = c_1 = 1 \qquad \& \qquad \dot{Q}(0) = -c_1 + \sqrt{5}c_2 = 1$$

You can solve these equations to obtain $c_1 = 1$ and $c_2 = 2/\sqrt{5}$. Therefore,

$$\boxed{Q(t) = e^{-t}\cos(\sqrt{5}\,t) + \frac{2}{\sqrt{5}}e^{-t}\sin(\sqrt{5}\,t)}.$$

Differentiate to find the current,

$$\boxed{I(t) = e^{-t}\Big(-\cos(\sqrt{5}\,t) - \sqrt{5}\sin(\sqrt{5}\,t)\Big) + \frac{2}{\sqrt{5}}e^{-t}\Big(-\sin(\sqrt{5}\,t) + \sqrt{5}\cos(\sqrt{5}\,t)\Big)}.$$

Example 3.9.8. Problem: the LC circuit or simple harmonic oscillator: *Suppose $R = 0$ and L, C are nonzero. If the circuit has current $I(0) = I_o$ and charge $Q(0) = Q_o$ then find the resulting equaitons for charge and current at time t.*

Solution: *We are faced with $L\ddot{Q} + \frac{1}{C}Q = 0$. This gives chararcteristic equation $\lambda^2 + \frac{1}{LC} = 0$ hence $\lambda = \pm i\sqrt{\frac{1}{LC}}$ and the general solution is, using $\omega = \sqrt{\frac{1}{LC}}$,*

$$Q(t) = c_1\cos(\omega t) + c_2\sin(\omega\,t)$$

Note that

$$\dot{Q}(t) = -c_1\omega\sin(\omega\,t) + c_2\omega\cos(\omega\,t).$$

Apply the given initial conditions,

$$Q(0) = c_1 = Q_o \qquad \& \qquad \dot{Q}(0) = c_2\omega = I_o$$

Therefore,

$$\boxed{Q(t) = Q_o\cos(\omega t) + \frac{I_o}{\omega}\sin(\omega\,t)}.$$

Differentiate to find the current,

$$\boxed{I(t) = -\omega Q_o\sin(\omega t) + I_o\cos(\omega\,t)}.$$

3.9.3 springs with external force

Suppose a mass m undergoes one-dimensional motion under the influence of a spring force $F_s = -kx$ and a velocity dependent friction force $F_f = -\beta\dot{x}$ **and** some external force f. Newton's Second Law states $m\ddot{x} = -kx - \beta\dot{x} + f$. We find

$$\boxed{m\ddot{x} + \beta\dot{x} + kx = f}$$

We have tools to solve this problem for many interesting forces.

Example 3.9.9. Problem: constant force: *Suppose $m \neq 0$ and $\beta, k > 0$. Suppose a constant force $f = F_o$ is placed on the spring. Describe the resulting motion.*

Solution: *We are faced with $m\ddot{x} + \beta\dot{x} + kx = F_o$. Notice that we can find x_h to solve the homogeneous, force-free equation; $m\ddot{x_h} + \beta\dot{x_h} + kx_h = 0$. The particular solution is simply $x_p = F_o/k$ and it follows the general solution has the form:*

$$x(t) = x_h(t) + F_o/k$$

We find motion that is almost identical to the problem with F_o removed. If we change coordinates to $y = x - F_o/k$ then clearly $\dot{x} = \dot{y}$ and $\ddot{x} = \ddot{y}$ hence $m\ddot{y} + \beta\dot{y} + ky = 0$. An important example of a constant force is that of gravity on a srping hanging vertically. The net-effect of gravity is to reset the equilbrium position of the spring from $x = 0$ to $x = mg/k$. The frequency of any oscillations is not effected by gravity, moreover, the spring returns to the new equilbrium $x = mg/k$ in the same manner as it would with matching damping, mass and stiffness in a horizontal set-up. For example, to find the frequency of oscillation for shocks on a car is determined from the viscosity of the oil in the shock assembly, the stiffness of the springs and the mass of the car. Gravity doesn't enter the picture.

Example 3.9.10. Problem: sinusoidal, nonresonant, force on a simple harmonic oscillator *Suppose $m = 1$ and $\beta = 0$ and $k = 1$. Suppose a sinusoidal force $f = F_o\cos(2t)$ is placed on the spring. Find the equations of motion given that $x(0) = 0$ and $\dot{x}(0) = 0$.*

Solution: *observe that $\ddot{x} + x = F_o\cos(2t)$ has homogeneous solution $x_h(t) = c_1\cos(t) + c_2\sin(t)$ and the method of annihilators can be used to indicate $x_p = A\cos(2t) + B\sin(2t)$. Calculate $\ddot{x}_p = -4x_p$ thus*

$$\ddot{x}_p + x_p = F_o\cos(2t) \quad \Rightarrow \quad -3A\cos(2t) - 3B\sin(2t) = F_o\cos(2t)$$

Thus $A = -F_o/3$ and $B = 0$ which gives us the general solution,

$$x(t) = c_1\cos(t) + c_2\sin(t) - \frac{F_o}{3}\cos(2t)$$

We calculate $\dot{x}(t) = -c_1\sin(t) + c_2\cos(t) + \frac{2F_o}{3}\sin(2t)$. Apply initial conditions to the solution,

$$c_1 - \frac{F_o}{3} = 0 \qquad \& c_2 = 0 \quad \Rightarrow \quad \boxed{x(t) = \tfrac{F_o}{3}\big[\cos(t) - \cos(2t)\big]}.$$

Example 3.9.11. Problem: sinsoidal, resonant, force on a simple harmonic oscillator: *Suppose $m = 1$ and $\beta = 0$ and $k = 1$. Suppose a sinusoidal force $f = 2\cos(t)$ is placed on the spring. Find the equations of motion given that $x(0) = 1$ and $\dot{x}(0) = 0$.*

Solution: *observe that $\ddot{x}+x = 2\cos(t)$ has homogeneous solution $x_h(t) = c_1\cos(t)+c_2\sin(t)$ and the method of annihilators can be used to indicate $x_p = At\cos(t) + Bt\sin(t)$. We calculate,*

$$\dot{x}_p = (A + Bt)\cos(t) + (B - At)\sin(t)$$

$$\ddot{x}_p = B\cos(t) - (A + Bt)\sin(t) - A\sin(t) + (B - At)\cos(t) = (2B - At)\cos(t) - (2A + Bt)\sin(t)$$

Now plug these into $\ddot{x}_p + x_p = 2\cos(t)$ to obtain:

$$At\cos(t) + Bt\sin(t) + (2B - At)\cos(t) - (2A + Bt)\sin(t) = 2\cos(t)$$

notice the terms with coefficients t cancel and we deduce $2B = 2$ and $-2A = 0$ thus $A = 0$ and $B = 1$. We find the general solution

$$x(t) = c_1\cos(t) + c_2\sin(t) + t\sin(t)$$

Note $\dot{x}(t) = -c_1\sin(t) + c_2\cos(t) + \sin(t) + t\cos(t)$. Apply the initial conditions, $x(0) = c_1 = 1$ and $\dot{x}(0) = c_2 = 0$. Therefore, the equation of motion is

$$\boxed{x(t) = \cos(t) + t\sin(t)}.$$

Note that as $t \to \infty$ the equation above ceases to be physically reasonable. In the absence of damping it is possible for the energy injected from the external force to just build and build leading to infinite energy. Of course the spring cannot store infinite energy and it breaks. In this case without damping it is simple enough to judge the absence or presence of resonance. Resonance occurs iff the forcing function has the same frequency as the natural frequency $\omega = \sqrt{\frac{k}{m}}$. In the case that there is damping we say **resonance** is reached if for a given m, β, k the applied force $F_o \cos(\gamma t)$ produces a particular solution of largest magnitude.

To keep it simple let us consider a damped spring in the arbitrary *underdamped* case where $\beta^2 - 4mk < 0$ with an external force $F_o \cos(\gamma t)$. We seek to study solutions of

$$m\ddot{x} + \beta\dot{x} + kx = F_o \cos(\gamma t)$$

Observe the characteristic equation is $m\lambda^2 + \beta\lambda + k = 0$ gives $\lambda^2 + \frac{\beta}{m}\lambda + \frac{k}{m} = 0$. Complete the square, or use the quadratic formula, whichever you prefer:

$$\lambda = \frac{-\beta \pm \sqrt{\beta^2 - 4mk}}{2m} = \frac{-\beta \pm i\sqrt{4mk - \beta^2}}{2m}$$

It follows that the homogeneous (also called the **transient** solution since it goes away for $t >> 0$) is

$$x_h(t) = e^{\frac{-\beta t}{2m}} \big(c_1 \cos(\omega t) + c_2 \sin(\omega t)\big)$$

where I defined $\omega = \frac{\sqrt{4mk-\beta^2}}{2m}$ for convenience. The particular solution is also called the **steady-state** solution since it tends to dominate for $t >> 0$. Suppose $x_p = A\cos(\gamma t) + B\sin(\gamma t)$ calculate,

$$\dot{x}_p = -\gamma A \sin(\gamma t) + \gamma B \cos(\gamma t) \qquad \& \qquad \ddot{x}_p = -\gamma^2 A \cos(\gamma t) - \gamma^2 B \sin(\gamma t)$$

Subtitute into $m\ddot{x}_p + \beta\dot{x}_p + kx_p = F_o \cos(\gamma t)$ and find

$$-m\gamma^2 A \cos(\gamma t) - m\gamma^2 B \sin(\gamma t) - \beta\gamma A \sin(\gamma t) + \beta\gamma B \cos(\gamma t) + kA\cos(\gamma t) + kB \sin(\gamma t) = F_o \cos(\gamma t)$$

Hence,

$$\big[-m\gamma^2 A + \beta\gamma B + kA\big] \cos(\gamma t) + \big[-m\gamma^2 B - \beta\gamma A + kB\big] \sin(\gamma t) = F_o \cos(\gamma t)$$

Equating coefficients yield the conditions:

$$(k - m\gamma^2)A + \beta\gamma B = F_o \qquad \& \qquad (k - m\gamma^2)B - \beta\gamma A = 0$$

We solve the second equation for $B = \frac{\beta\gamma}{k - m\gamma^2} A$ and substitute this into the other equation,

$$(k - m\gamma^2)A + \frac{\beta^2\gamma^2}{k - m\gamma^2} A = F_o$$

Now make a common denominator,

$$\frac{(k - m\gamma^2)^2 + \beta^2\gamma^2}{k - m\gamma^2} A = F_o$$

We find,

$$A = \frac{(k - m\gamma^2)F_o}{(k - m\gamma^2)^2 + \beta^2\gamma^2} \qquad \& \qquad B = \frac{\beta\gamma F_o}{(k - m\gamma^2)^2 + \beta^2\gamma^2}$$

It follows that the particular solution has the form

$$x_p = \frac{F_o}{(k - m\gamma^2)^2 + \beta^2\gamma^2} \Big[(k - m\gamma^2)\cos(\gamma t) + \beta\gamma \sin(\gamma t)\Big]$$

You can show[14] that the amplitude of $A_1\cos(\gamma t) + A_2\sin(\gamma t)$ is given by $A = \sqrt{A_1^2 + A_2^2}$. Apply this lemma to the formula above to write the particular solution in the simplified form

$$\boxed{x_p = \frac{F_o}{\sqrt{(k - m\gamma^2)^2 + \beta^2\gamma^2}}\sin(\gamma t + \phi)}$$

where ϕ is a particular angle. We're mostly interested in the magnitude so let us focus our attention on the amplitude of the steady state solution[15].

Suppose k, m, β are fixed and let us study $M(\gamma) = \frac{1}{\sqrt{(k-m\gamma^2)^2+\beta^2\gamma^2}}$. What choice of γ maximizes this factor thus producing the resonant motion? Differentiate and seek the critical value:

$$\frac{dM}{d\gamma} = -\frac{1}{2}\cdot\frac{2(k - m\gamma^2)(-2m\gamma) + 2\beta^2\gamma}{[(k - m\gamma^2)^2 + \beta^2\gamma^2]^{3/2}} = 0$$

The critical value must arise from the vanishing of the numerator since the denominator is nonzero,

$$(k - m\gamma^2)(-2m\gamma) + \beta^2\gamma = 0 \quad \Rightarrow \quad (-2mk + 2m^2\gamma^2 + \beta^2)\gamma = 0$$

But, we already know $\gamma = 0$ is not the frequency we're looking for, thus

$$-2mk + 2m^2\gamma^2 + \beta^2 = 0 \quad \Rightarrow \quad \gamma = \pm\sqrt{\frac{2mk - \beta^2}{2m^2}}$$

Nothing is lost by choosing the + here and we can simplify to find

$$\boxed{\gamma_c = \sqrt{\frac{k}{m} - \frac{\beta^2}{2m^2}}}$$

It is nice to see that $\beta = 0$ returns us to the natural frequency $\omega = \sqrt{\frac{k}{m}}$ as we studied initially. Section 4.10 of Nagel Saff and Snider, or 6-3 of Ritger & Rose if you would like to see further analysis.

3.10 RLC circuit with a voltage source

Suppose a resistor R, an inductor L and a capacitor C are placed in series with a voltage source $\mathcal{E}$. Kirchoff's Voltage Law reads

$$\boxed{L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = \mathcal{E}}$$

We can solve these problems in the same way as we have just explored for the spring force problem. I will jump straight to the resonance problem and change gears a bit to once more promote complex notation.

Suppose we have an underdamped R, L, C circuit driven by a voltage source $\mathcal{E}(t) = V_o\cos(\gamma t)$. I propose we solve the related complex problem

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = V_o e^{i\gamma t}$$

We propose a complex particular solution: $Q_p = Ae^{i\gamma t}$ hence

$$Q_p' = i\gamma Ae^{i\gamma t} \qquad \& \qquad Q_p'' = -\gamma^2 Ae^{i\gamma t}$$

[14]the precalculus chapter in my calculus I notes has some of the ideas needed for this derivation

[15]see page 240-241 of Nagel Saff and Snider for a few comments beyond mine and a nice picture to see the difference between the transient and steady state solutions

Subtitute into $LQ_p'' + RQ_p' + \frac{1}{C}Q_p = V_oe^{i\gamma t}$ and factor out the imaginary exponential

$$\big[-\gamma^2 L + i\gamma R + \frac{1}{C}\big]Ae^{i\gamma t} = V_oe^{i\gamma t}$$

Hence,

$$-\gamma^2 L + i\gamma R + \frac{1}{C} = \frac{V_o}{A}$$

Hence,

$$A = \frac{V_o}{1/C - \gamma^2 L + i\gamma R} \cdot \frac{1/C - \gamma^2 L - i\gamma R}{1/C - \gamma^2 L - i\gamma R} = \frac{V_o[1/C - \gamma^2 L - i\gamma R]}{(1/C - \gamma^2 L)^2 + \gamma^2 R^2}$$

Thus, using $e^{i\gamma t} = \cos(\gamma t) + i\sin(\gamma t)$, the complex particular solution is given by

$$Q_p(t) = \left[\frac{V_o[1/C - \gamma^2 L]}{(1/C - \gamma^2 L)^2 + \gamma^2 R^2} - i\frac{V_o\gamma R}{(1/C - \gamma^2 L)^2 + \gamma^2 R^2}\right]\left[\cos(\gamma t) + i\sin(\gamma t)\right].$$

We can read solutions for particular solutions of any real linear combination of $V_o\cos(\gamma t)$ and $V_o\sin(\gamma t)$. For example, for $L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = V_o\cos(\gamma t)$ we derive the particular solution

$$Q_{p_1}(t) = \frac{V_o[1/C - \gamma^2 L]}{(1/C - \gamma^2 L)^2 + \gamma^2 R^2}\cos(\gamma t) + \frac{V_o\gamma R}{(1/C - \gamma^2 L)^2 + \gamma^2 R^2}\sin(\gamma t)$$

Likewise, as $Im(e^{i\gamma t}) = \sin(\gamma t)$ the solution of $L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = V_o\sin(\gamma t)$ is given by the $Im(Q_p) = Q_{p_2}$.

$$Q_{p_2}(t) = \frac{V_o[1/C - \gamma^2 L]}{(1/C - \gamma^2 L)^2 + \gamma^2 R^2}\sin(\gamma t) + \frac{V_o\gamma R}{(1/C - \gamma^2 L)^2 + \gamma^2 R^2}\cos(\gamma t)$$

To solve $L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = B_1V_o\cos(\gamma t) + B_2V_o\sin(\gamma t)$ we use superposition to form the particular solution $Q_{p_3} = B_1Q_{p_1} + B_2Q_{p_2}$.

Remark 3.10.1.

Notice that Q_{p_1} is analagous to the solution we found studying resonance for the underdamped spring. If we use the dictionary $m \leftrightarrows L$, $\beta \leftrightarrows R$, $k \leftrightarrows 1/C$, $F_o \leftrightarrows V_o$ then it ought to be obvious the solution above was already derived in real notation. However, the complex solution is quicker and cleaner. We also can deduce that resonance is reached at

$$\gamma_r = \sqrt{\frac{k}{m} - \frac{\beta^2}{2m^2}} \leftrightarrows \sqrt{\frac{1}{LC} - \frac{R^2}{2L^2}}$$

and note how $R = 0$ reduces the problem to the pure harmonic oscillation of the LC-tank.

3.11 practice problems

PP 73 Solve $y'' - y' - 11y = 0$ where $y' = dy/dx$.

PP 74 Solve $4w'' + 20w' + 25w = 0$ where $w' = dw/dx$.

PP 75 Solve $y'' + 2y' + y = 0$ where $y(0) = 1$ and $y'(0) = -3$ given $y' = dy/dx$.

PP 76 Suppose $y_1 = te^{2t}$ and $y_2 = e^{2t}$. Determine if y_1 and y_2 are linearly dependent on $(0, 1)$.

PP 77 Solve $y'' - 8y' + 7y = 0$ where $y = y(t)$.

PP 78 Solve $z'' + 10z' + 25z = 0$ where $z' = dz/dx$

PP 79 Solve $u'' + 7u = 0$ given t is the independent variable.

PP 80 Solve $y'' + 10y' + 41y = 0$ given $y' = dy/dx$.

PP 81 Solve $y'' - 2y' + 2y = 0$ given $y(\pi) = e^{\pi}$ and $y'(\pi) = 0$. Use independent variable x.

PP 82 Consider $y'' - 6y' - 4y = 4\sin(3t) - t^2e^{3t} + \frac{1}{t}$. Can we solve this via the method of undetermined coefficients ? If so, suggest a form for the particular solution.

PP 83 Consider $y'' - 2y' + 3y = \cosh t$. Can we solve this via the method of undetermined coefficients ? If so, suggest a form for the particular solution.

PP 84 Find the general solution to $y'' - y = 1 - 11t$.

PP 85 Solve $z'' + z = 2e^{-x}$ given $z(0) = 0$ and $z'(0) = 0$.

PP 86 Determine if $\{\sin^2 x, \cos^2 x, 1\}$ is linearly independent on $\mathbb{R}$.

PP 87 Show $\{x, x^2x^3, x^4\}$ is linearly independent on $\mathbb{R}$.

PP 88 Let us define $L[y] = y''' + y' + xy$. Let $y_1 = \sin x$ and $y_2 = x$.

- **(a.)** Calculate $L[y_1]$ and $L[y_2]$,
- **(b.)** Solve $L[y] = 2x\sin x - x^2 - 1$,
- **(c.)** Solve $L[y] = 4x^2 + 4 - 6x\sin x$.

I am not asking for the general solution in the problem above

PP 89 Solve $y''' + 2y'' - 8y' = 0$ given $y' = dy/dt$.

PP 90 Solve $u''' - 9u'' + 27u' - 27u = 0$ given $u = u(x)$.

PP 91 Solve $y^{(4)} + 4y'' + 4y = 0$ given $y = y(x)$.

PP 92 Solve $y^{(4)} + 2y''' + 10y'' + 18y' + 9y = 0$ given that $y = \sin(3x)$ is a solution.

PP 93 Let $D = d/dx$. Solve

$$(D + 1)^2(D - 6)^3(D + 5)(D^2 + 1)(D^2 + 4)[y] = 0.$$

PP 94 Completely factor the following polynomials over $\mathbb{R}$. Place any irreducible quadratic factors in the completed-square format $(x - \alpha)^2 + \beta^2$.

- **(a.)** $x^2 + 6x + 20$
- **(b.)** $x^4 + 5x^2 - 6$
- **(c.)** $x^4 - 256$
- **(d.)** $f(x) = -20 - 36x - 15x^2 + 5x^3 + 5x^4 + x^5$ given that $f(-1) = 0$ and $f(-2 + i) = 0$

PP 95 Find the general solutions of the DEqns given below.

- **(a.)** $y'' + 6y' + 20y = 0$
- **(b.)** $(D^4 + 5D^2 - 6)[y] = 0$
- **(c.)** $y^{(4)} - 256y = 0$
- **(d.)** $-20y - 36y' - 15y'' + 5y''' + 5y^{(4)} + y^{(5)} = 0$
  given that $y_1 = e^{-x}$ and $y_2 = e^{-2x}\cos(x)$ are solutions.

PP 96 Solve the following ODE,

$$(D^2 + 6D + 13)(D^2 - 9)(D^2 + 1)(D^2 + 4D + 3)[y] = 0.$$

PP 97 Find minimal annihilators for each of the functions below:

- **(a.)** $f_1(x) = x^2e^x$

(b.) $f_2(x) = e^x \cos(4x)$

(c.) $f_3(x) = x^3 + e^x \cos(4x)$

(d.) $f_4(x) = \cos^2(3x) + e^x \cosh(x)$

PP 98 Set-up, but do not determine explicitly, the particular solutions for:

(a.) $y'' - 2y' + y = x^2 e^x$

(b.) $y'' + 16y = e^x \cos(4x)$

(c.) $y''' + y' = x^3 + e^x \cos(4x)$

(d.) $y''' + 36y' = \cos^2(3x) + e^x \cosh(x)$

PP 99 Solve $y'' + 3y' + 2y = x + e^{-x} + e^{3x}$.

PP 100 Solve $y'' + 3y' + 2y = e^{-2t}\cos(t)$.

PP 101 Solve $y'' + 3y' + 2y = 20(t + e^{-t} + e^{3t}) + 2e^{-2t}\cos(t)$ given that $y(0) = 0$ and $y'(0) = 1$.

PP 102 Solve $y'' + 2y' + y = \frac{e^{-x}}{x+1}$.

PP 103 Solve $y'' + y = \tan^2(x)$

PP 104 Find integral solutions for $y''' + 16y' = f$. (you need to use variation of parameters, I would explicitly calculate the determinants of S_1, S_2, S_3 as I discuss in the notes)

PP 105 Solve $x^2 y'' - (x^2 + 2x)y' + (x + 2)y = x^3$. Note $y_1 = x$ is a fundamental solution of the DEqn. *Hint: find the 2nd. fundamental soln. and then use variation of parameters to find* y_p*...*

PP 106 Solve the following cauchy euler problems. Give your solution as a real linear combination of the real-value functions in the fundamental solution set.

(a.) $4x^2 y'' + y = 0$

(b.) $x^2 y'' - 3xy' + 5y = 0$

(c.) $2x^2 y'' + 3xy' - y = 0$

(d.) $x^3 y''' + 2x^2 y'' - xy' + y = 0$

(e.) $x^2 y'' + 5xy' + 4y = 0$ with $y(1) = 2$ and $y'(1) = -3$

PP 107 Derive a formula to rewrite $x^4 D^4$ as a polynomial in xD. Use the result to solve $x^4 D^4[y] = 0$. Please use my notes for formulas for $x^3 D^3$ and $x^2 D^2$, also, use Leibniz product rule for best results.

PP 108 Suppose a mass of 1kg is attached to a spring with stiffness 5 Newtons per meter. Then the spring and mass are immersed in an oil with viscosity producing a velocity-dependent friction force with coefficient $\beta = 4kg/s$. If a force $F(t) = 10\cos(t)$ (in Newtons and seconds) is used to drive the system then what is the resulting equation of motion? Assume that $x(0) = 0$ and $v(0) = 1$. What anglular frequency γ would make the force $10\cos(\gamma t)$ give a particular solution of maximum amplitude?

PP 109 Suppose an RLC-circuit is assembled with $R = 11\Omega$, $L = 1H$ and $C = 0.1F$. If a half-decaying voltage source of $\mathcal{E}(t) = 10e^{-t} + \cos(t)$ is attached to the circuit then what is the resulting current as a function of time. Assume a switch closes at $t = 0$ connecting the voltage source to the circuit. This means $I(0) = 0$ and $Q(0) = 0$.

PP 110 Let f and g be functions which are twice continuously differentiable on an interval I for which $W(f, g; x) \neq 0$ for each $x \in I$. Show that

$$det \begin{bmatrix} y & y' & y'' \\ f & f' & f'' \\ g & g' & g'' \end{bmatrix} = 0$$

is a second order, linear, homogeneous differential equation with fundamental solutions $y_1 = f$ and $y_2 = g$. Then, use this result to construct a differential equation which has solutions e^x and $e^{1/x}$, include the interval on which these are the solutions.

PP 111 Show that the Cauchy-Euler problem

$$a_o \frac{d^n y}{x^n} + a_1 \frac{d^{n-1} y}{x^{n-1}} + \cdots + a_{n-1} \frac{dy}{dx} + a_n y = 0$$

problem changes to a constant coefficient problem if we make the substitution $x = e^t$. Use this result to derive the solutions of the Cauchy-Euler problem for which we find $R = 1$ three times, or $R = 1+2i$ twice.

PP 112 Novel methods of integration.

(a.) Solve $\int x^3 e^x \, dx$ by solving $\frac{dy}{dx} = x^3 e^x$ using the method of undetermined coefficients.

(b.) Solve $\int e^x \cos(2x) \, dx$ by studying the integral of $\int e^{(1+2i)x} \, dx$. *Hint: we know* $\frac{d}{dx} e^{\lambda x} = \lambda e^{\lambda x}$ *even for the case* $\lambda = 1 + 2i$.

PP 113 Let y_1 and y_2 form the fundamental solution set of the second order linear differential equation

$$a_o y'' + a_1 y' + a_2 y = 0$$

on an interval I. Show that between any two successive zeros of y_1 there is exactly one zero of y_2.

PP 114 (Ritger & Rose section 5-4 problem 1a-d) find the general solution of:

(a.) $y'' = 0$

(b.) $y'' - 2y' = 0$

(c.) $y'' - a^2 y = 0$

(d.) $y'' + a^2 y = 0$

PP 115 (Ritger & Rose section 5-4 problem 3) Suppose $ay'' + by' + cy = 0$ has distinct real characteristic values of $\lambda_\pm = A \pm B$ and hence a general solution $y = c_1 e^{\lambda_+ x} + c_2 e^{\lambda_- x}$. Show that the general solution can be rewritten as

$$y = e^{Ax}(b_1 \cosh(Bx) + b_2 \sinh(Bx)).$$

PP 116 (Ritger & Rose section 5-5 problems 1,2,3 and 8)

(1.) $y'' + 3y' - 5y = 4e^{2x} + 6e^{-3x}$

(2.) $y'' + 3y' + 5y = 2\sin(3x)$

(3.) $y'' + 9y = 4\cos(3x)$

(8.) $y'' - 3y' = 2x^2 + 3e^x$

PP 117 (introduction to theory of adjoints, from page 95 of Boyce and DiPrima's 3rd Ed.) If

$$p(x)y'' + q(x)y' + r(x)y = 0$$

can be expressed as $[p(x)y']' + [f(x)y]' = 0$ then it is said to be **exact**. Omit x-dependence in p, q, r, μ for brevity, if $py'' + qy' + ry = 0$ is not exact then it is possible to make it exact with multiplication by the appropriate integrating factor μ. **Show** that for μ to accomplish its stated task it must itself be the solution of the so-called **adjoint equation**

$$p\mu'' + (2p' - q)\mu' + (p'' - q' + r)\mu = 0.$$

where we have assumed p, q possess the stated derivatives. Find the adjoint equation for

(a.) *constant coefficient case:* $ay'' + by' + cy = 0$

(b.) *Bessel Eqn. of order* ν: $x^2 y'' + xy' + (x^2 - \nu^2)y = 0$

(c.) *The Airy Equation:* $y'' - xy = 0$

PP 118 Consider the differential equation $y''' - 3y'' + 2y' = g(t)$. Is $\{1, e^t, e^{2t}\}$ a fundamental solution set ? Explain your answer.

PP 119 Let $y_1(x) = x^3$ and $y_2(x) = |x|^3$. Show that $W(y_1, y_2)(x) = 0$ for all $x \in \mathbb{R}$. However, explain why $\{y_1, y_2\}$ is linearly independent on $\mathbb{R}$. Does there exist a linear ODE for which $\{y_1, y_2\}$ forms the fundamental solution set? Discuss.

PP 120 Solve

(a.) $y'' + 5y' + 6y = 0$,

(b.) $y'' + 4y' + 4y = 0$,

(c.) $y'' + 4y' + 5y = 0$.

PP 121 Solve

(a.) $y'' - 36y = 0$ subject the initial conditions $y(0) = 1, y'(0) = 0$,

(b.) $y'' + 25 = 0$ subject the initial conditions $y(0) = 1, y'(0) = 0$.

PP 122 Solve, here $D = d/dx$

(a.) $D^2(D^2 - 9)[y] = 0$,

(b.) $(D^2 + 6D + 18)^2[y] = 0$,

(c.) $(D^2 + 3D + 2)(D^2 - 4)[y] = 0$

PP 123 Give constant coefficient ODEs for which the following form general solutions. Please leave your answer in $D = d/dx$ factored notation. No need to multiply them out.

(a.) $y = c_1 e^{-4x} + c_2 e^{-3x}$,

(b.) $y = c_1 e^{10x} + c_2 x e^{10x}$

(c.) $y = A\cosh(3x + B)$

(d.) $y = c_1 + Ae^{2x}\sin(3x + \phi)$

PP 124 (fitting initial conditions) Given $x(t) = c_1 \cos\omega t + c_2 \sin\omega t$ is the general solution to

$$x'' + \omega^2 x = 0.$$

Show $x(0) = x_o$ and $x'(0) = x_1$ implies $c_1 = x_o$ and $c_2 = x_1/\omega$.

PP 125 (reduction of order) Use the reduction of order formula $y_2 = y_1 \int \frac{exp(-\int pdx)}{y_1^2} dx$ to calculate a second linearly independent solution for $x^2 + 2xy' - 6y = 0$ given $y_1 = x^2$.

PP 126 (reduction of order) Consider $x^2y'' - 3xy' + 5y = 0$ for $x > 0$. You are given that $y_1 = x^2\cos\ln x$ is a solution. Find y_2 for which y_1, y_2 forms a fundamental solution set for the given differential equation. One approach is to use the $n = 2$ reduction of order formula as derived in 3.6 of my notes.

PP 127 (based on Cook section 3.7) Suppose $T = D$ and $S = 3 - x^2D$. Solve

(a.) $ST[y] = 0$,

(b.) $TS[y] = 0$.

PP 128 Consider $f(x) = x^{2+3i}$ for $x > 0$. Find u, v such that $f = u + iv$. Furthermore, by differentiation of u, v show that $f'(x) = (2 + 3i)x^{1-3i}$.
(the point: you can replace 2 with $a \in \mathbb{R}$ and 3 with $b \in \mathbb{R}$ and derive

$$\frac{d}{dx} x^{a+ib} = (a + ib)x^{a-1+ib};$$

we see the power rule extends naturally to the case of a complex exponent of the power function. This is an important fact as we deal with solving the Cauchy Euler problem $ax^2y'' + bxy' + cy = 0$)

PP 129 Solve the following Cauchy Euler problems

(a.) $4x^2y'' + y = 0$,

(b.) $25x^2y'' + 25xy' + y = 0$

(c.) $x^3y''' - 6y = 0$

PP 130 Suppose that y_1 is a nontrivial solution of $y'' + p(x)y' + q(x)y = 0$. We seek a method to derive a second LI solution. Let y_2 be such a solution and **show that it must satisfy**

$$\frac{d}{dx}\left[\frac{y_2}{y_1}\right] = \frac{W(y_1, y_2)}{y_1^2}.$$

Now, use Abel's formula to find a nice formula for y_2.

PP 131 (from page 103 of Boyce and DiPrima's 3rd Ed.) Consider for $N \in \mathbb{N}$,

$$xy'' - (x + N)y' + Ny = 0.$$

(a.) show $y_1 = e^x$ is a solution.

(b.) show that $y_2 = ce^x \int x^N e^{-x}\, dx$ is a second solution. (perhaps use the result of the previous problem, or the theorem from my notes or Ritger & Rose)

(c.) set $c = \frac{-1}{N!}$ and show by induction that $y_2(x) = T_n(x)$ where $T_n(x)$ denotes the n-th order Taylor polynomial of e^x centered at zero.

PP 132 Find minimal annihilators for each of the functions below:

(a.) $f_1(x) = x^2e^x$

(b.) $f_2(x) = e^x \cos(4x)$

(c.) $f_3(x) = x^3 + e^x \cos(4x)$

(d.) $f_4(x) = \cos^2(3x) + e^x \cosh(x)$

Now, given what you've just thought through, set-up, but do not determine explicitly, the particular solutions for:

(a.) $y'' - 2y' + y = x^2e^x$

(b.) $y'' + 16y = e^x \cos(4x)$

(c.) $y''' + y' = x^3 + e^x \cos(4x)$

(d.) $y''' + 36y' = \cos^2(3x) + e^x \cosh(x)$

PP 133 (Zill section 4.4 problem 17) Solve $y'' - 2y' + 5y = e^x \cos(2x)$.

PP 134 Solve $y'' + 3y' + 2y = t^2$ subject the initial conditions $y(0) = 1$ and $y'(0) = 0$.

PP 135 (Zill section 4.5 problem 63) Solve $y^{(4)} - 2y''' + y'' = e^x + 1$.

PP 136 (Zill section 4.6 problem 8) Solve $y'' - y = \sinh(2x)$.

PP 137 (Zill section 4.6 problem 14) Solve $y'' - 2y' + y = e^t \tan^{-1}(t)$.

PP 138 Solve $y'' + 3y' + 2y = x + e^{-x} + e^{3x}$.

PP 139 Solve $y'' + 2y' + y = \frac{e^{-x}}{x+1}$.

PP 140 Find integral solutions for $y''' + 16y' = f$. (you need to use variation of parameters, I would explicitly calculate the determinants of S_1, S_2, S_3 as I discuss in the notes)

PP 141 Solve $x^2y'' - (x^2 + 2x)y' + (x + 2)y = x^3$. Note $y_1 = x$ is a fundamental solution of the DEqn. *Hint: find the 2nd. fundamental soln. and then use variation of parameters to find y_p...*

PP 142 Solve $(xD+3)(D^2-4)[y]=0$. Be careful.

PP 143 Suppose L is a linear differential operator. Furthermore, suppose $L[y_1]=g_1$ and $L[y_2]=2g_1$. Solve $L[y]=0$ using the given solutions.

PP 144 Find an integral solution for $y''+y=g$ with $y(0)=y_o$ and $y'(0)=y_1$ and g is some integrable function of time t.

PP 145 Consider a third order linear differential equation for which $\sin(x), \cos(x)$ and $\ln(x)$ appear as the fundamental solution set. Call this differential equation $L[y]=0$. Solve $L[y]=42$ via variation of parameters. It is interesting to note that even though I asked you to supply an explicit linear ODE $L[y]=0$ to solve you should not need that explicit formula to solve $L[y]=42$.

PP 146 Green's function for a linear ODE $L[y]=f$ provides a method for solving the DEqn via integration. If we assume the initial conditions of the given ODE are all trivial then the operator L can be inverted; $L[y]=f$ with trivial initial conditions iff $y=L^{-1}[f]$. In particular, if $G(x,t)$ is a function for which $y(x)=\int_{x_o}^{x} G(x,t)f(t)dt$ is a solution of $L[y]=f$ then we say G is a **Green's function** for L.

In the case of a second order differential equation with fundamental solutions y_1, y_2 (with $L[y_1]=0$ and $L[y_2]=0$ for LI y_1, y_2) we can construct a Green's function as follows:

$$G(x,t)=\frac{y_1(t)y_2(x)-y_1(x)y_2(t)}{y_1(t)y_2'(t)-y_2(t)y_1'(t)}$$

Then observe $y=\int_{x_o}^{x} G(x,t)f(t)dt$ gives a solution to $L[y]=f$ by variation of parameters. Find Green's function for the following solution sets and write an integral solution for $L[y]=f$ for the given L and given initial conditions:

(a.) $L=D^2+9$, $y_1=\cos 3t$ and $y_2=\sin 3t$ with $y(0)=y'(0)=0$,

(b.) $L=D^2+3D+2$, $y_1=e^{-x}$, $y_2=e^{-2x}$ with $y(0)=-1$ and $y'(0)=0$,

PP 147 Use the Green's function technique to solve

$$y''+3y'+2y=\sin(e^x)$$

subject $y(0)=-1$ and $y'(0)=0$. In other words, work out the integrals for part (b.) of the previous problem given that $f(x)=\sin(e^x)$.

PP 148 Suppose a spring is attached to a mass of 1 kg and the spring has spring constant 16 N/m. This spring mass system is immersed in an oil which gives a retarding frictional force of $F_{retard}=-\beta v$ where v is velocity and $\beta=10\ Ns/m$. Find the equations of motion (please omit units, so in the usual notation $m=1$, $k=16$ and $\beta=10$) in the cases

(a.) $x(0)=-1$ and $x'(0)=0$

(b.) $x(0)=-1$ and $x'(0)=12$

PP 149 Newton's Law for a retarded spring-mass system with external force f yield

$$m\ddot{x}+\beta\dot{x}+kx=f$$

Given $m=2$, $b=0$, $k=32$ and $f=68e^{-2t}\cos(4t)$ find the equation of motion given the system has initial conditions $x(0)=\dot{x}(0)=0$.

PP 150 Consider Newton's Second Law for mass-spring system under a sinusoidal force:

$$\ddot{x}+\omega^2x=F_o\cos\gamma t$$

given $x(0)=\dot{x}(0)=0$. Here $F_o, \omega\gamma$ are nonzero constants.

(a.) Find $x(t)$ given that $\gamma \neq \omega$

(b.) Calculate $x_r(t) = \lim_{\gamma \to \omega} x(t)$

(c.) Contrast the motion of $x(t)$ and $x_r(t)$ as $t \to \infty$

PP 151 Kirchoff's Voltage Law for an RLC-circuit with voltage source $\mathcal{E}$ is given by

$$L\frac{dI}{dt} + RI + \frac{1}{C}Q = \mathcal{E}$$

Since $I = \frac{dQ}{dt}$ we find $L\ddot{Q} + R\dot{Q} + Q/C = \mathcal{E}$. Given that $L = 1$ and $R = 2$ and $C = 0.25$ and $\mathcal{E} = 50\cos t$ find the charge Q as a function of time t given the initial charge and current are both zero for $t = 0$.

PP 152 If we study the motion of an spring

$$m\ddot{x} + \beta\dot{x} + kx = F$$

such that $\beta^2 - 4mk < 0$ then it is known as **underdamped motion**. If the external force $F = F_o\cos(\gamma t)$ then we find the motion is dominated by the particular solution as $t \to \infty$. Let $\omega = \frac{\sqrt{4mk-\beta^2}}{2m}$, then the homogeneous solution $x_h(t) = e^{\frac{-\beta t}{2m}}\left(c_1\cos(\omega t) + c_2\sin(\omega t)\right) \to 0$ as $t \to \infty$. Show that the particular solution of such a system is given by

$$x_p = \frac{F_o\sin(\gamma t + \phi)}{\sqrt{(k - m\gamma^2)^2 + \beta^2\gamma^2}}$$

where ϕ is a constant. Then, find the frequency γ which maximizes the magnitude of x_p in the following cases:

(a.) $m = 1/2$ and $k = 19$ and $\beta = 1$

(b.) $m = 1$ and $k = 2$ and $\beta = \sqrt{6}$.

PP 153 Find general solution of $y'' - 3y' + 2y = 0$ where $y' = dy/dx$.

PP 154 Find general solution of $y'' - 6y' + 9y = 0$ where $y' = dy/dx$.

PP 155 Find general solution of $y'' + 6y' + 13y = 0$ where $y' = dy/dt$.

PP 156 Find general solution of $(D-2)^3(D^2-1)D^2[y] = 0$ where $D = d/dx$.

PP 157 Suppose $D = d/dx$ and $L = D^n + a_{n-1}D^{n-1} + \cdots + a_2D^2 + a_1D + a_o$ defines differential equation $L[y] = 0$. Find smallest n and the coefficients $a_o, a_1, \dots, a_{n-1} \in \mathbb{R}$ for which $e^x\cos(2x)$ and x^3 are solutions to the differential equation $L[y] = 0$.

PP 158 Find general solution of $y'' - 9y = t^2 + e^t + 1$.

PP 159 Solve $y'' + y = 2\cos t + \sin t$

PP 160 Solve $y'' + 4y = \tan(2x)$.

PP 161 Find general solution of $y'' + 3y' + 2y = t + 1$.

PP 162 Solve $y'' + y = \cos t + e^t$

PP 163 Solve $y'' - 6y + 9y = 0$ where $y' = dy/dt$.

PP 164 Solve $((D+3)^2 + 36)[y] = 0$ where $D = d/d\theta$.

PP 165 Let $D = d/dx$. Observe $(D^4 + 9D^2)[y] = x + \cos(x)$ can be solved by the method of undetermined coefficients aided by the annihilator method. We find the minimal particular solution derived from the annhilator method is: (circle one answer)

(a.) $y_p = Ax + B + C\cos(x) + D\sin(x)$

(b.) $y_p = Ax^3 + Bx^2 + Cx\cos(x) + Dx\sin(x)$

(c.) $y_p = Ax^3 + Bx^2 + C\cos(x) + D\sin(x)$

(d.) $y_p = A + Bx + C\cos(3x) + D\sin(3x)$

PP 166 A spring has mass $m = 1$, coefficient of damping $\beta = 4$ and a spring constant $k = 5$. Find the general solution of Newton's Second Law.

PP 167 Solve $y' - 3y = 2x + 3$.

PP 168 Find a particular solution for $y'' - 4y' + 3y = 65\cos(2t)$.

PP 169 Find a particular solution of $y'' - 4y' + 3y = e^t$.

PP 170 Find the general solution of $y'' - 4y' + 3y = 130\cos(2t) + 7e^t$.

PP 171 Suppose $L[y] = 0$ is an n-th order differential equation where $L = D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_o$ and $D = d/dt$ and $a_{n-1}, \dots, a_1, a_o \in \mathbb{R}$. If $L[\, e^t\cos(2t)\,] = 0$ and $L[\, t^3 e^{-t}\,] = 0$ then find the smallest n which allows these solutions and give the explicit form of L in terms of $D = d/dt$. You need not multiply out the formula, I am perfectly happy with L in factored form.

PP 172 Find an integral solution for $x > 0$ to the Cauchy Euler problem
$x^2y'' + xy' + 9y = g$ where g is a continuous function.

PP 173 Solve the following differential equations:

(a.) $y'' - 8y' + 7y = 0$ where $y' = dy/dt$,

(b.) $z'' + 10z + 25z = 0$ where $z' = dz/dx$,

(c.) $u'' + 7u = 0$ where $u' = du/dt$,

(d.) $y'' + 10y' + 41y = 0$ where $y' = dyd/dx$

(e.) $y''' + 4y'' + 5y' = 0$ where $y' = dy/dx$.

PP 174 Find the minimal annihilator for each of the following functions: for each define D as either $D = d/dx$ or $D = d/dt$ as appropriate:

(a.) $g = e^x + \sin(4x)$

(b.) $g = x^2 + \cosh(x)$

(c.) $g = te^{-3t} + 2$

(d.) $g = \cos(x)\sin(3x)$

(e.) $g = e^t\cos(6t)$

PP 175 Set-up, but do not explicitly determine the coefficients, the form of y_p via the method of annihilators. Notice you found the annihilators in the previous problem.

(a.) $y'' - y = e^x + \sin(4x)$

(b.) $y'' + y' = x^2 + \cosh(x)$

(c.) $y'' + 3y' = te^{-3t} + 2$

(d.) $y'' + 4y = \cos(x)\sin(3x)$

(e.) $y'' + 36y = e^t\cos(6t)$

PP 176 Solve $y'' - 4y' = 6t + e^t$.

PP 177 Solve $y'' + 2y' + y = \cos(x) + 3$ subject the initial conditions $y(0) = 0$ and $y'(0) = 1$.

PP 178 Consider $mx'' + bx' + kx = 0$ where $m > 0$ and $b, k \geq 0$. Show that in every possible case the motion of the solution is bounded.

PP 179 Find the general solution of

(a.) $y'' + \ y = 3\cos(2x)$

(b.) $y'' + \ y = \csc(x)$

(c.) $y'' + \ y = 2\csc(x) + \cos(2x)$

PP 180 These require variation of parameters technique.

(a.) Solve $y'' - 2y' + y = \frac{1}{t}e^t$

(b.) Solve $y'' + y = \sec^3\theta$.

PP 181 Solve the integral $\int (x^3 + 2x)e^x \, dx = y$ by solving $\frac{dy}{dx} = (x^2 + 2x)e^x$ via the method of undetermined coefficients

PP 182 Consider the differential equation given by: $D = d/dx$ and

$$(D^4 + 2D^3 + 10D^2 + 18D + 9)[y] = 0$$

You are given that $y = \sin 3x$ is a solution to the above. Use this data to help solve the problem.

Chapter 4

systems of ordinary differential equations

A system of differential equations is a collection of differential equations which we wish to solve simultaneously. This is the same terminology as we use for a system of linear eqautions. Indeed, the systems we consider in this work are formed from constant coefficient ODEs. This means we can use constant matrices to formulate our system of differential equations. Moreover, since **reduction of order** always allows us to recast a given problem as a system of first order ODEs, we simply face the problem of the form

$$\frac{d\vec{r}}{dt} = A\vec{r} + \vec{f}$$

where A and $\vec{f}$ are given and $\vec{r} = \langle x_1, x_2, \dots, x_n \rangle$ is a vector of unknown functions we wish to calculate. The theory here is almost identical to our previous work on the n-th order linear ODE. Once again we need to construct n-LI solutions and find a particular solution $\vec{r}_p$ in order to formulate the general solution:

$$\vec{r} = c_1\vec{r}_1 + c_2\vec{r}_2 + \cdots + c_n\vec{r}_n + \vec{r}_p.$$

Notice the differences though: n is now the number of separate first order differential equations which form our system and the solutions are vector-valued functions of time. That said, we do form the **fundamental solution set** via linearly independent solutions $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n$ of the corresponding homogeneous system

$$\frac{d\vec{r}}{dt} = A\vec{r}.$$

It turns out the construction of the fundamental solution set requires we master some elmementary linear algebra, in particular the algebra of **eigenvectors**. As before, we face a characeristic equations whose roots reveal much about the possible form of the solution set. But, in contrast to the n-th order problem, there is structure which is not revealed by the characteristic equation alone. Sometimes we find enough eigenvectors or complex eigenvectors to form the whole solution set, but not always. This distinction is intimately tied to the problem of diagonalization of matrices and we will spend a little time and energy to appreciate how the linear algebra we need for solving systems also allows us to solve algebra problems which require **uncoupling**. Applications of the algebra we study extend well past our application to systems of ODEs. For example, with this algebra we can understand the second derivative test for multivariate functions of two variables (which we teach in Calculus III) and higher dimensional analogs which are not known to standard Calculus III courses. For another example, we can discover the principle axes for the inertia tensor which tells us natural directions for a rigid object to spin. The applications of diagonalization are vast.

To complete our study of systems of ODEs we need deeper linear algebra and the **matrix exponential**.

$$e^{tA} = I + tA + \frac{1}{2}t^2A^2 + \frac{1}{3!}t^3A^3 + \cdots$$

In particular, we need chains of generalized eigenvectors and even chains of generalized complex eigenvectors. These chains unlock the power of the magic formula:

$$e^{tA} = e^{\lambda t}\left(I + t(A-\lambda I) + \frac{t^2}{2}(A-\lambda I)^2 + \frac{t^3}{3!}(A-\lambda I)^3 + \cdots \right).$$

In particular, the formula above brings multiples of t and t^2 where appropriate. The matrix exponenital is a **fundamental solution matrix**, however, it does require a particular set of skills to apply with success. It turns out that the matrix expoential is central to understanding the continuous symmetries of modern physics. Ask me about spinors and matrix groups sometime if you're curious.

Finally, once we have settled the construction of the fundamental set of (homogeneous) solutions we turn to study the nonhomogeneous problem. We'll see how variation of parameters allows a simple construction of the particular solution based on integration. We focus on the 2×2 case since inversion of larger symbolic matrices requires signinficantly more calculation.

We also share the theory of linear systems including the study of linear independence and existence and uniqueness theorems. Once more the existence of a fundamental solution set informs how to construct the general solution. The theory covers much more than we study in this chapter. Systems with variable coefficients are not covered here. We may cover some such systems with the method of Laplace transforms later in this course.

4.1 calculus and matrices

The construction of matrices and the operations thereof are designed to simplify arguments about algebraic systems of linear equations. We will see that the matrix is also of great utitility for the solution of systems of linear differential equations. We've already seen how matrix calculations unify and simplify with the theory of the Wronskian and the technique of variation of parameters. I now pause to introduce and define explicitly the algebra and construction of matrices and we also derive some important theorems about their calculus.

A $p \times q$ **matrix** over R is an ordered array of pq-objects from R which has p-**rows** and q-**columns**. The objects in the matrix are called its **components**. In particular, if matrix A has components $A_{ij} \in R$ for i, j with $1 \leq i \leq p$ and $1 \leq j \leq q$ then we denote the array by:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \cdots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{bmatrix} = [A_{ij}]$$

We also view a matrix as columns or rows glued together:

$$A = \big[col_1(A)|col_2(A)|\cdots|col_q(A)\big] = \begin{bmatrix} row_1(A) \\ \hline row_2(A) \\ \hline \vdots \\ \hline row_p(A) \end{bmatrix}$$

where we define $col_j(A) = [A_{1j}, A_{2j}, \dots, A_{pj}]^T$ and $row_i(A) = [A_{i1}, A_{i2}, \dots, A_{iq}]$. The set of all $p \times q$ matrices assembled from objects in R is denoted $R^{p \times q}$. Notice that if $A, B \in R^{p \times q}$ then $A = B$ iff $A_{ij} = B_{ij}$ for all i, j with $1 \leq i \leq p$ and $1 \leq j \leq q$. In other words, two matrices are **equal** iff all the matching components are equal. We use this principle in many definitions, for example: if $A \in R^{p \times q}$ then the **transpose** $A^T \in R^{q \times p}$ is defined by $A^T_{ij} = A_{ji}$ for all i, j.

We are primarily interested in the cases $R = \mathbb{R}, \mathbb{C}$ or some suitable set of functions. All of these spaces allow for addition and multiplication of the components. It is therefore logical to define the sum, difference, scalar multiple and product of matrices as follows:

Definition 4.1.1. *If $A, B \in R^{p\times q}$ and $C \in R^{q\times r}$ and $c \in R$ then define*

$$(A+B)_{ij} = A_{ij} + B_{ij} \qquad (A-B)_{ij} = A_{ij} - B_{ij} \qquad (cA)_{ij} = cA_{ij} \qquad (BC)_{ik} = \sum_{j=1}^{q} B_{ik}C_{kj}.$$

This means that $(A+B), (A-B), cA \in R^{p\times q}$ whereas $BC \in R^{p\times r}$. The **matrix product** of a $p \times q$ and $q \times r$ matrix is a $p \times r$ matrix. In order for the product BC to be defined we must have the rows in B be the same size as the columns in C. We can express the product in terms of dot-products:

$$\boxed{(BC)_{ik} = row_i(B) \bullet col_k(C)}$$

Let me give a few examples to help you understand these formulas.

Example 4.1.2. *The product of a 3×2 and 2×3 is a 3×3*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} [1,0]\bullet[4,7] & [1,0]\bullet[5,8] & [1,0]\bullet[6,9] \\ [0,1]\bullet[4,7] & [0,1]\bullet[5,8] & [0,1]\bullet[6,9] \\ [0,0]\bullet[4,7] & [0,0]\bullet[5,8] & [0,0]\bullet[6,9] \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \\ 0 & 0 & 0 \end{bmatrix}$$

Example 4.1.3. *The product of a 3×1 and 1×3 is a 3×3*

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 4\cdot 1 & 5\cdot 1 & 6\cdot 1 \\ 4\cdot 2 & 5\cdot 2 & 6\cdot 2 \\ 4\cdot 3 & 5\cdot 3 & 6\cdot 3 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{bmatrix}$$

Example 4.1.4. *Let $A = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right]$. We calculate*

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \\ &= \begin{bmatrix} [1,2]\bullet[5,7] & [1,2]\bullet[6,8] \\ [3,4]\bullet[5,7] & [3,4]\bullet[6,8] \end{bmatrix} \\ &= \begin{bmatrix} 5+14 & 6+16 \\ 15+28 & 18+32 \end{bmatrix} \\ &= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \end{aligned}$$

Notice the product of square matrices is square. For numbers $a, b \in \mathbb{R}$ it we know the product of a and b is commutative ($ab = ba$). Let's calculate the product of A and B in the opposite order,

$$\begin{aligned} BA &= \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} [5,6]\bullet[1,3] & [5,6]\bullet[2,4] \\ [7,8]\bullet[1,3] & [7,8]\bullet[2,4] \end{bmatrix} \\ &= \begin{bmatrix} 5+18 & 10+24 \\ 7+24 & 14+32 \end{bmatrix} \\ &= \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \end{aligned}$$

Clearly $AB \neq BA$ thus matrix multiplication is **noncommutative** *or* **nonabelian**.

When we say that matrix multiplication is noncommuative that indicates that the product of two matrices does not *generally* commute. However, there are special matrices which commute with other matrices.

Example 4.1.5. *Let* $I = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ *and* $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. *We calculate*

$$IA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Likewise calculate,

$$AI = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Since the matrix A *was arbitrary we conclude that* $IA = AI$ *for all* $A \in \mathbb{R}^{2\times 2}$.

The Kronecker delta δ_{ij} is defined to be zero if $i \neq j$ and $\delta_{ii} = 1$. The **identity matrix** is the matrix I such that $I_{ij} = \delta_{ij}$. It is simple to show that $AI = A$ and $IA = A$ for all matrices.

Definition 4.1.6.

Let $A \in \mathbb{R}^{\,n\times n}$. If there exists $B \in \mathbb{R}^{\,n\times n}$ such that $AB = I$ and $BA = I$ then we say that A is **invertible** and $A^{-1} = B$. Invertible matrices are also called **nonsingular**. If a matrix has no inverse then it is called a **noninvertible** or **singular** matrix.

Example 4.1.7. *In the case of a* 2×2 *matrix* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *a nice formula to find the inverse is known provided* $det(A) = ad - bc \neq 0$:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

It's not hard to show this formula works,

$$\begin{aligned} \tfrac{1}{ad-bc}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} &= \tfrac{1}{ad-bc}\begin{bmatrix} ad-bc & -ab+ab \\ cd-dc & -bc+da \end{bmatrix} \\ &= \tfrac{1}{ad-bc}\begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

This formula is worth memorizing for future use.

The problem of inverting an $n \times n$ matrix is more challenging when $n > 2$. However, it is generally true[1] that A^{-1} exists iff $det(A) \neq 0$. Recall our discussion of Cramer's rule in the variation of parameters section, we divided by the determinant for form the solution. If the determinant is zero then we cannot use Cramer's rule and we must seek other methods of solution. In particular, the methods of Gaussian elimination or back subsitution are general and we will need to use those techniques to solve the eigenvector problem in the later part of this chapter. But, don't let me get too ahead of the story. Let's finish our tour of matrix algebra.

Proposition 4.1.8.

If A, B are invertible square matrices and c is nonzero then

1. $(AB)^{-1} = B^{-1}A^{-1}$,
2. $(cA)^{-1} = \frac{1}{c}A^{-1}$,

[1] the formula is simply $A^{-1} = \frac{1}{det(A)}ad(A)$ where $ad(A)$ is the adjoint of A, see my linear notes where I give the explicit calculation for an arbitrary 3×3 case

Proof: property (1.) is called the **socks-shoes** property because in the same way you first put on your socks and then your shoes to invert the process you first take off your shoes then your socks. The proof is just a calculation:

$$(AB)B^{-1}A^{-1} = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I.$$

The proof of (2.) is similar □

The power of a matrix is defined in the natural way. Notice we need for A to be square in order for the product AA to be defined.

Definition 4.1.9.

Let $A \in \mathbb{R}^{\ n\times n}$. We define $A^0 = I$, $A^1 = A$ and $A^m = AA^{m-1}$ for all $m \geq 1$. If A is invertible then $A^{-p} = (A^{-1})^p$.

As you would expect, $A^3 = AA^2 = AAA$.

Proposition 4.1.10.

Let $A, B \in \mathbb{R}^{\ n\times n}$ and $p, q \in \mathbb{N} \cup \{0\}$

1. $(A^p)^q = A^{pq}$.
2. $A^pA^q = A^{p+q}$.
3. If A is invertible, $(A^{-1})^{-1} = A$.

You should notice that $(AB)^p \neq A^pB^p$ for matrices. Instead,

$$(AB)^2 = ABAB, \qquad (AB)^3 = ABABAB, etc...$$

This means the binomial theorem will not hold for matrices. For example,

$$(A + B)^2 = (A + B)(A + B) = A(A + B) + B(A + B) = AA + AB + BA + BB$$

hence $(A+B)^2 \neq A^2 +2AB +B^2$ as the matrix product is not generally commutative. If we have A and B commute then $AB = BA$ and we can prove that $(AB)^p = A^pB^p$ and the binomial theorem holds true.

Example 4.1.11. *A square matrix A is said to be* **idempotent** *of order k if there exists $k \in \mathbb{N}$ such that $A^{k-1} \neq I$ and $A^k = I$. On the other hand, a square matrix B is said to be* **nilpotent** *of order k if there exists $k \in \mathbb{N}$ such that $B^{k-1} \neq 0$ and $B^k = 0$. Suppose B is nilpotent of order 2; $B^2 = 0$ and $B \neq 0$. Let $X = I + B$ and calculate,*

$$X^2 = (I + B)(I + B) = II + IB + BI + B^2 = I + 2B$$

$$X^3 = (I + B)(I + 2B) = II + I2B + BI + B2B = I + 3B$$

You can show by induction that $X^k = I + kB$. (neat, that is all I have to say for now)

Example 4.1.12. *A square matrix which only has zero entries in all components except possibly the diagonal is called a* **diagonal matrix**. *We say $D \in R^{n\times n}$ is diagonal iff $D_{ij} = 0$ for $i \neq j$. Consider, if $X = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}$ and $Y = \begin{bmatrix} y_1 & 0 \\ 0 & y_2 \end{bmatrix}$ then we find*

$$XY = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}\begin{bmatrix} y_1 & 0 \\ 0 & y_2 \end{bmatrix} = \begin{bmatrix} x_1y_1 & 0 \\ 0 & x_2y_2 \end{bmatrix} = \begin{bmatrix} y_1x_1 & 0 \\ 0 & y_2x_2 \end{bmatrix} = YX.$$

These results extend beyond the 2×2 case. If X, Y are diagonal $n \times n$ matrices then $XY = YX$. You can also show that if X is diagonal and A is any other square matrix then $AX = XA$. We will later need the formula below:

$$\begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & D_n \end{bmatrix}^k = \begin{bmatrix} D_1^k & 0 & \cdots & 0 \\ 0 & D_2^k & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & D_n^k \end{bmatrix}.$$

Example 4.1.13. *The product of a* 2×2 *and* 2×1 *is a* 2×1*. Let* $A = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ *and let* $v = \left[\begin{smallmatrix} 5 \\ 7 \end{smallmatrix}\right]$*,*

$$Av = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} [1,2] \bullet [5,7] \\ [3,4] \bullet [5,7] \end{bmatrix} = \begin{bmatrix} 19 \\ 43 \end{bmatrix}$$

Likewise, define $w = \left[\begin{smallmatrix} 6 \\ 8 \end{smallmatrix}\right]$ *and calculate*

$$Aw = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} [1,2] \bullet [6,8] \\ [3,4] \bullet [6,8] \end{bmatrix} = \begin{bmatrix} 22 \\ 50 \end{bmatrix}$$

Something interesting to observe here, recall that in Example 4.1.4 we calculated $AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$*. But these are the same numbers we just found from the two matrix-vector products calculated above. We identify that* B *is just the* **concatenation** *of the vectors* v *and* w*;* $B = [v|w] = \left[\begin{array}{c|c} 5 & 6 \\ 7 & 8 \end{array}\right]$*. Observe that:*

$$AB = A[v|w] = [Av|Aw].$$

The term **concatenate** is sometimes replaced with the word **adjoin**. I think of the process as gluing matrices together. This is an important operation since it allows us to lump together many solutions into a single matrix of solutions.

Proposition 4.1.14.

Let $A \in \mathbb{R}^{\,m \times n}$ and $B \in \mathbb{R}^{\,n \times p}$ then we can understand the matrix multiplication of A and B as the concatenation of several matrix-vector products,

$$AB = A[col_1(B)|col_2(B)|\cdots|col_p(B)] = [Acol_1(B)|Acol_2(B)|\cdots|Acol_p(B)]$$

The proof is left to the reader. Finally, to conclude our brief tour of matrix algebra, I collect all my favorite properties for matrix multiplication in the theorem below. To summarize, matrix math works as you would expect with the exception that matrix multiplication is not commutative. We must be careful about the order of letters in matrix expressions.

Example 4.1.15. *Suppose* $Ax = b$ *has solution* x_1 *and* $Ax = c$ *has solution* x_2 *then note that* $X_o = [x_1|x_2]$ *is a* **solution matrix** *of the matrix equation* $AX = [b|c]$*. In particular, observe:*

$$AX_o = A[x_1|x_2] = [Ax_1|Ax_2] = [b|c].$$

For the sake of completeness and perhaps to satisfy the curiousity of the inquisitive student I pause to give a brief synopsis of how we solve systems of equations with matrix techniques. We will not need technology to solve most problems we confront, but I think it is useful to be aware of just how you can use the ”rref” command to solve any linear system.

Remark 4.1.16. *summary of how to solve linear equations*

(1.) Write the system of equations in matrix notation $Au = b$

(2.) Perform Gaussian eliminate to reduce the augmented coefficient matrx $[A|b]$ to its reduced-row echelon form $rref[A|b]$ (usually I use a computer for complicated examples)

(3.) Read the solution from $rref[A|b]$. There are three cases:

   (a.) there are no solutions

   (b.) there is a unique solution

   (c.) there are infinitely many solutions

The nuts and bolts of gaussian elimination is the process of adding, subtracting and multiplying equations by a nonzero constant towards the goal of eliminating as many variables as possible.
Let us illustrate the remark above.

Example 4.1.17. *Suppose* $u_1+u_2 = 3$ *and* $u_1-u_2 = -1$. *Then* $Au = b$ *for coefficient matrix* $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ *and* $b = [3, -1]^T$. *By gaussian elimination,*

$$rref \left[\begin{array}{cc|c} 1 & 1 & 3 \\ 1 & -1 & -1 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right]$$

It follows that $u_1 = 1$ *and* $u_2 = 2$. *This is the* **unique solution**. *The solution set* $\{(1, 2)\}$ *contains a single solution.*

Set aside matrix techniques, you can solve the system above by adding equations to obtain $2u_1 = 2$ hence $u_1 = 1$ and $u_2 = 3 - 1 = 2$.

Example 4.1.18. *Suppose* $u_1 + u_2 + u_3 = 1$ *and* $2u_1 + 2u_2 + 2u_3 = 4$. *Then* $Au = b$ *for coefficient matrix* $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$ *and* $b = [1, 4]^T$. *By gaussian elimination,*

$$rref \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 4 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

The second row suggests that $0u_1 + 0u_2 + 0u_3 = 1$ *or* $0 = 1$ *which is clearly false hence the system is* **inconsistent** *and the solution set in this case is the empty set.*

Set aside matrix techniques, you can solve the system above by dividing the second equation by 2 to reveal $u_1 + u_2 + u_3 = 2$. Thus insisting both equations are simultaneously true amounts to insisting that $1 = 2$. For this reason the system has no solutions.

Example 4.1.19. *Suppose* $u_1 + u_2 + u_3 = 0$ *and* $2u_1 + 2u_2 + 2u_3 = 0$. *Then* $Au = b$ *for coefficient matrix* $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$ *and* $b = [0, 0]^T$. *By gaussian elimination,*

$$rref \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 2 & 2 & 0 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The second row suggests that $0u_1 + 0u_2 + 0u_3 = 0$ *or* $0 = 0$ *which is clearly true. This system is* **consistent** *and the solutions have* $u_1 + u_2 + u_3 = 0$. *It follows that the solution set is infinite*

$$\{[-u_2 - u_3, u_2, u_3]^T \mid u_2, u_3 \in \mathbb{R}\}.$$

Any solution can be written as $u_2[-1, 1, 0]^T + u_3[-1, 0, 1]^T$ *for particular constants* u_2, u_3.

It turns out that the last example is the type of matrix algebra problem we wil face with the eigenvector method. The theorem that follows summarizes the algebra of matrices.

Theorem 4.1.20.

If $A, B, C \in \mathbb{R}^{\ m\times n}$, $X, Y \in \mathbb{R}^{\ n\times p}$, $Z \in \mathbb{R}^{\ p\times q}$ and $c_1, c_2 \in \mathbb{R}$ then

(1.) $(A+B)+C = A+(B+C)$,

(2.) $(AX)Z = A(XZ)$,

(3.) $A+B = B+A$,

(4.) $c_1(A+B) = c_1A + c_2B$,

(5.) $(c_1+c_2)A = c_1A + c_2A$,

(6.) $(c_1c_2)A = c_1(c_2A)$,

(7.) $(c_1A)X = c_1(AX) = A(c_1X) = (AX)c_1$,

(8.) $1A = A$,

(9.) $I_mA = A = AI_n$,

(10.) $A(X+Y) = AX + AY$,

(11.) $A(c_1X + c_2Y) = c_1AX + c_2AY$,

(12.) $(A+B)X = AX + BX$,

The proof of the theorem above follows easily from the definitions of matrix operation. I give some explicit proof in my linear algebra notes. In fact, all of the examples thus far are all taken from my linear algebra notes where I discuss not just these formulas, but also their motivation from many avenues of logic. The example that follows would not be something I would commonly include in the linear algebra course.

Example 4.1.21. *Suppose $R = C^\infty(\mathbb{R})$ be the set of all smooth functions on $\mathbb{R}$. For example,*

$$A = \begin{bmatrix} \cos(t) & t^3 \\ 3^t & \ln(t^2+1) \end{bmatrix} \in R^{2\times 2}.$$

We can multiply A above by 3^{-t} by

$$3^{-t}A = \begin{bmatrix} 3^{-t}\cos(t) & t^3 3^{-t} \\ 1 & \ln(t^2+1)3^{-t} \end{bmatrix} \in R^{2\times 2}.$$

We can subtract the identity matrix to form $A - I$:

$$A - I = \begin{bmatrix} \cos(t) & t^3 \\ 3^t & \ln(t^2+1) \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(t)-1 & t^3 \\ 3^t & \ln(t^2+1)-1 \end{bmatrix}$$

Another way of looking at A in the example above is that it is a **matrix-valued** function of a real variable t. In other words, $A : \mathbb{R} \to \mathbb{R}^{2\times 2}$; this means for each $t \in \mathbb{R}$ we assign a single matrix $A(t) \in \mathbb{R}^{2\times 2}$. We can similarly consider $p \times q$-matrix valued functions of a real variable[2] . We now turn to the calculus of such matrices.

[2]for those of you who have (or are) taking linear algebra, the space $R^{p\times q}$ is not necessarily a vector space since R is not a field in some examples. The space of smooth functions forms what is called a ring and the set of matrices over a ring can be understood as a "module". A module is like a vector space where the scalar multiplication is taken from a ring rather than a field. Every vector space is a module but some modules are not vector spaces.

Definition 4.1.22.

A matrix-valued function of a real variable is a function from $I \subseteq \mathbb{R}$ to $\mathbb{R}^{m\times n}$. Suppose $A : I \subseteq \mathbb{R} \to \mathbb{R}^{m\times n}$ is such that $A_{ij} : I \subseteq \mathbb{R} \to \mathbb{R}$ is differentiable for each i, j then we define

$$\tfrac{dA}{dt} = \left[\tfrac{dA_{ij}}{dt}\right]$$

which can also be denoted $(A')_{ij} = A'_{ij}$. We likewise define $\int A dt = [\int A_{ij} dt]$ for A with integrable components. Definite integrals and higher derivatives are also defined component-wise.

Example 4.1.23. *Suppose* $A(t) = \left[\begin{array}{cc} 2t & 3t^2 \\ 4t^3 & 5t^4 \end{array}\right]$. *I'll calculate a few items just to illustrate the definition above. calculate; to differentiate a matrix we differentiate each component one at a time:*

$$A'(t) = \left[\begin{array}{cc} 2 & 6t \\ 12t^2 & 20t^3 \end{array}\right] \qquad A''(t) = \left[\begin{array}{cc} 0 & 6 \\ 24t & 60t^2 \end{array}\right] \qquad A'(0) = \left[\begin{array}{cc} 2 & 0 \\ 0 & 0 \end{array}\right]$$

Integrate by integrating each component:

$$\int A(t)dt = \left[\begin{array}{cc} t^2 + c_1 & t^3 + c_2 \\ t^4 + c_3 & t^5 + c_4 \end{array}\right] \qquad \int_0^2 A(t)dt = \left[\begin{array}{cc} t^2\big|_0^2 & t^3\big|_0^2 \\ & \\ t^4\big|_0^2 & t^5\big|_0^2 \end{array}\right] = \left[\begin{array}{cc} 4 & 8 \\ 16 & 32 \end{array}\right]$$

Example 4.1.24. *Suppose* $A = \left[\begin{array}{cc} t & 1 \\ 0 & t^2 \end{array}\right]$. *Calculate* $A^2 = \left[\begin{array}{cc} t & 1 \\ 0 & t^2 \end{array}\right]\left[\begin{array}{cc} t & 1 \\ 0 & t^2 \end{array}\right]$ *hence,*

$$A^2 = \left[\begin{array}{cc} t^2 & t + t^2 \\ 0 & t^4 \end{array}\right]$$

Clearly $\frac{d}{dt}[A^2] = \left[\begin{array}{cc} 2t & 1 + 2t \\ 0 & 4t^3 \end{array}\right]$. *On the other hand, calculate*

$$2A\frac{dA}{dt} = 2\left[\begin{array}{cc} t & 1 \\ 0 & t^2 \end{array}\right]\left[\begin{array}{cc} 1 & 0 \\ 0 & 2t \end{array}\right] = 2\left[\begin{array}{cc} t & 2t \\ 0 & 2t^3 \end{array}\right] = \left[\begin{array}{cc} 2t & 4t \\ 0 & 4t^3 \end{array}\right] \neq \frac{d}{dt}[A^2]$$

The naive chain-rule fails.

Theorem 4.1.25.

Suppose A, B are matrix-valued functions of a real variable, f is a function of a real variable, c is a constant, and C is a constant matrix then

1. $(AB)' = A'B + AB'$ (product rule for matrices)
2. $(AC)' = A'C$
3. $(CA)' = CA'$
4. $(fA)' = f'A + fA'$
5. $(cA)' = cA'$
6. $(A + B)' = A' + B'$

where each of the functions is evaluated at the same time t and I assume that the functions and matrices are differentiable at that value of t and of course the matrices A, B, C are such that the multiplications are well-defined.

Proof: Suppose $A(t) \in \mathbb{R}^{\ m\times n}$ and $B(t) \in \mathbb{R}^{\ n\times p}$ consider,

$$\begin{aligned}
(AB)'_{ij} &= \tfrac{d}{dt}((AB)_{ij}) && \text{defn. derivative of matrix} \\
&= \tfrac{d}{dt}(\textstyle\sum_k A_{ik}B_{kj}) && \text{defn. of matrix multiplication} \\
&= \textstyle\sum_k \tfrac{d}{dt}(A_{ik}B_{kj}) && \text{linearity of derivative} \\
&= \textstyle\sum_k \left[\tfrac{dA_{ik}}{dt}B_{kj} + A_{ik}\tfrac{dB_{kj}}{dt}\right] && \text{ordinary product rules} \\
&= \textstyle\sum_k \tfrac{dA_{ik}}{dt}B_{kj} + \sum_k A_{ik}\tfrac{dB_{kj}}{dt} && \text{algebra} \\
&= (A'B)_{ij} + (AB')_{ij} && \text{defn. of matrix multiplication} \\
&= (A'B + AB')_{ij} && \text{defn. matrix addition}
\end{aligned}$$

this proves (1.) as i, j were arbitrary in the calculation above. The proof of (2.) and (3.) follow quickly from (1.) since C constant means $C' = 0$. Proof of (4.) is similar to (1.):

$$\begin{aligned}
(fA)'_{ij} &= \tfrac{d}{dt}((fA)_{ij}) && \text{defn. derivative of matrix} \\
&= \tfrac{d}{dt}(fA_{ij}) && \text{defn. of scalar multiplication} \\
&= \tfrac{df}{dt}A_{ij} + f\tfrac{dA_{ij}}{dt} && \text{ordinary product rule} \\
&= (\tfrac{df}{dt}A + f\tfrac{dA}{dt})_{ij} && \text{defn. matrix addition.}
\end{aligned}$$

The proof of (5.) follows from taking $f(t) = c$ which has $f' = 0$. I leave the proof of (6.) as an exercise for the reader. $\square$.

To summarize: the calculus of matrices is the same as the calculus of functions with the small qualifier that we must respect the rules of matrix algebra. The noncommutativity of matrix multiplication is the main distinguishing feature.

Let's investigate, just for the sake of some practice mostly, what the non-naive chain rule for the square of matrix function.

Example 4.1.26. *Let $A : \mathbb{R} \to \mathbb{R}^{\ n\times n}$ be a square-matrix-valued differentiable function of a real variable t. Calculate, use the product rule:*

$$\frac{d}{dt}[A^2] = \frac{d}{dt}[AA] = \frac{dA}{dt}A + A\frac{dA}{dt}.$$

In retrospect, it must be the case that the matrix A does not commute with $\frac{dA}{dt}$ in Example 4.1.24. The noncommutative nature of the matrix multiplication is the source of the naive chain-rule not working in the current context. In contrast, we have seen that the chain-rule for complex-valued functions of a real variable does often work. For example, $\frac{d}{dt}e^{\lambda t} = \lambda e^{\lambda t}$ or $\frac{d}{dx}x^{\lambda} = \lambda x^{\lambda-1}$. It is possible to show that if $f(z)$ is analytic and $g(t)$ is differentiable from $\mathbb{R}$ to $\mathbb{C}$ then $\frac{d}{dt}f(g(t)) = \frac{df}{dz}(g(t))\frac{dg}{dt}$ where $\frac{df}{dz}$ is the derivative of f with respect to the complex variable z. However, you probably will not discuss this in complex variables since it's not terribly interesting in the big-scheme of that course. I find it interesting to contrast to the matrix case here. You might wonder if there is a concept of differentiation with respect to a matrix, or differentiation with respect to a vector. The answer is yes. However, I leave that for some other course.

Example 4.1.27. *Another example for fun. The set of* **orthogonal matrices** *is denoted $O(n)$ and is defined to be the set of $n \times n$ matrices A such that $A^TA = I$. These matrices correspond to changes of coordinate which do not change the length of vectors; $A\vec{x} \bullet A\vec{y} = \vec{x} \bullet \vec{y}$. It turns out that $O(n)$ is made from rotations and relections.*

Suppose we have a curve of orthogonal matrices; $A : \mathbb{R} \to O(n)$ then we know that $A^T(t)A(t) = I$ for all $t \in \mathbb{R}$. If the component functions are differentiable then we can differentiate this equation to learn about the structure that the tangent vector to an orthogonal matrix must possess. Observe:

$$\frac{d}{dt}[A^T(t)A(t)] = \frac{d}{dt}[I] \quad \Rightarrow \quad \frac{dA}{dt}^T A(t) + A^T(t)\frac{dA}{dt} = 0$$

Suppose the curve we considered passed through the identity matrix I (which is in $O(n)$ as $I^TI = I$) and suppose this happened at $t = 0$ then we have

$$\frac{dA^T}{dt}(0) + \frac{dA}{dt}(0) = 0$$

Let $B = \frac{dA}{dt}(0)$ then we see that $B^T = -B$ is a necessary condition for tangent vectors to the orthogonal matrices at the identity matrix. A matrix with $B^T = -B$ is said to be **antisymmetric** *or* **skew-symmetric**. *The space of all such skew matrices is called $o(n)$. The set $O(n)$ paired with matrix multiplication is called a* **Lie Group** *whereas the set $o(n)$ paired with the matrix commutator is called a* **Lie Algebra**[3]. *These are concepts of considerable interest in modern studies of differential equations.*

4.2 Row Reduction Technique for Solving Systems

A system of three equations and three unknowns x_1, x_2, x_3 is given by:

$$\begin{array}{c} A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = b_1 \\ A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = b_2 \\ A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = b_3 \end{array} \Rightarrow \underbrace{\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}}_{b} \tag{4.1}$$

Notice we can express three scalar equations as one vector equation $Ax = b$. We call A the **coefficient matrix** and b is the **inhomogeneous term**. The easiest notation to work out the solution of $Ax = b$ is the **augmented coefficient matrix** $[A|b]$ which frees us from writing x_1, x_2, x_3. I should mention, when $b = 0$ the system is called **homogeneous**. Notice, when the system is homogeneous there is always the zero solution since $A(0) = 0$.

We now take a little detour to discuss the standard method to solve linear systems. I should emphasize, this is not usually needed for the linear systems we solve in service of solving an system of ODEs. In fact, often those systems are far more easily solved by educated guessing guided by a few salient observations. Therefore, I may not present this material in lecture. I put it here for the students who may never take a proper course in linear algebra.

To begin we need to identify three basic operations we do when solving systems of equations. I'll define them for system of 3 equations and 3 unknowns, but it should be obvious this generalizes to m equations and n unknowns without much thought. The following operations are called **Elementary Row Operations**.

(1.) scaling row 1 by nonzero constant c

$$\left[\begin{array}{ccc|c} A_{11} & A_{12} & A_{13} & b_1 \\ A_{21} & A_{22} & A_{23} & b_2 \\ A_{31} & A_{32} & A_{33} & b_3 \end{array}\right] \xrightarrow{cr_1} \left[\begin{array}{ccc|c} cA_{11} & cA_{12} & cA_{13} & cb_1 \\ A_{21} & A_{22} & A_{23} & b_2 \\ A_{31} & A_{32} & A_{33} & b_3 \end{array}\right]$$

(2.) replace row 1 with the sum of row 1 and row 2

$$\left[\begin{array}{ccc|c} A_{11} & A_{12} & A_{13} & b_1 \\ A_{21} & A_{22} & A_{23} & b_2 \\ A_{31} & A_{32} & A_{33} & b_3 \end{array}\right] \xrightarrow{r_1 + r_2} \left[\begin{array}{ccc|c} A_{11} + A_{21} & A_{12} + A_{22} & A_{13} + A_{23} & b_1 + b_2 \\ A_{21} & A_{22} & A_{23} & b_2 \\ A_{31} & A_{32} & A_{33} & b_3 \end{array}\right]$$

(3.) swap rows 1 and 2

$$\left[\begin{array}{ccc|c} A_{11} & A_{12} & A_{13} & b_1 \\ A_{21} & A_{22} & A_{23} & b_2 \\ A_{31} & A_{32} & A_{33} & b_3 \end{array}\right] \xrightarrow{r_1 \longleftrightarrow r_2} \left[\begin{array}{ccc|c} A_{21} & A_{22} & A_{23} & b_2 \\ A_{11} & A_{12} & A_{13} & b_1 \\ A_{31} & A_{32} & A_{33} & b_3 \end{array}\right]$$

[3]it's is pronounced "Lee" not as you might expect

Each of the operations above corresponds to an allowed operation on a system of linear equations. When we make these operations we will not change the solution set. The Gauss-Jordan algorithm tells us which order to make these operations in order to reduce the matrix to a particularly simple format called the "reduced row echelon form" (I abbreviate this rref most places).

Definition 4.2.1. Gauss-Jordan Algorithm.

Given an m by n matrix M the following sequence of steps is called the Gauss-Jordan algorithm or Gaussian elimination. I define terms such as **pivot column** and **pivot position** as they arise in the algorithm below.

Step 1: Determine the leftmost nonzero column. This is a **pivot column** and the topmost position in this column is a **pivot position**.

Step 2: Perform a row swap to bring a nonzero entry of the pivot column to the topmost row which does not already contain a pivot position. (in the first iteration this will be the top row of the matrix)

Step 3: Add multiples of the pivot row to create zeros **below** the pivot position. This is called "clearing out the entries below the pivot position".

Step 4: If there are no more nonzero rows below the last pivot row then go to step 5. Otherwise, there is a nonzero row below the pivot row and the new pivot column is the next nonzero column to the right of the old pivot column. Go to step 2.

Step 5: the leftmost entry in each nonzero row is called the **leading entry** (these are the entries in the pivot positions). Scale the bottommost nonzero row to make the leading entry 1 and use row additions to clear out any remaining nonzero entries **above** the leading entries.

Step 6: If step 5 was performed on the top row then stop, otherwise apply Step 5 to the next row up the matrix.

Steps (1.)-(4.) are called the **forward pass**. A matrix produced by a foward pass is called the $r$educed echelon form of the matrix and it is denoted ref(A). Steps (5.) and (6.) are called the **backwards pass**. The matrix produced by completing Steps (1.)-(6.) is called the $r$educed row echelon form of M and it is denoted $rref(M)$.

The $ref[A|b]$ is not unique because there may be multiple choices for how Step 2 is executed. On the other hand, it turns out that $rref[A|b]$ is unique. [4] The backwards pass takes the ambiguity out of the algorithm. Notice the forward pass goes down the matrix while the backwards pass goes up the matrix.

Example 4.2.2. *Consider the system of equations* $\begin{array}{c} x_1+2x_2-3x_3=1 \\ 2x_1+4x_2=7 \\ -x_1+3x_2+2x_3=0 \end{array}$ *This gives augmented coefficient matrix* $[A|b]=\left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 2 & 4 & 0 & 7 \\ -1 & 3 & 2 & 0 \end{array}\right]$ *and we calculate* $rref[A|b]$ *as follows:*

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 2 & 4 & 0 & 7 \\ -1 & 3 & 2 & 0 \end{array}\right] \xrightarrow{r_2-2r_1} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 0 & 0 & 6 & 5 \\ -1 & 3 & 2 & 0 \end{array}\right] \xrightarrow{r_3+r_1}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 0 & 0 & 6 & 5 \\ 0 & 5 & -1 & 1 \end{array}\right] \xrightarrow{r_2\leftrightarrow r_3} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 0 & 5 & -1 & 1 \\ 0 & 0 & 6 & 5 \end{array}\right] = ref[A|b]$$

[4] The proof of uniqueness can be found in Appendix E of the text *Elementary Linear Algebra: A Matrix Approach*, 2nd ed. by Spence, Insel and Friedberg.

that completes the forward pass. We begin the backwards pass,

$$ref[A|b] = \left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 0 & 5 & -1 & 1 \\ 0 & 0 & 6 & 5 \end{array}\right] \xrightarrow{\frac{1}{6}r_3} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 0 & 5 & -1 & 1 \\ 0 & 0 & 1 & 5/6 \end{array}\right] \xrightarrow{r_2 + r_3} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 1 \\ 0 & 5 & 0 & 11/6 \\ 0 & 0 & 1 & 5/6 \end{array}\right] \xrightarrow{r_1 + 3r_3}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 21/6 \\ 0 & 5 & 0 & 11/6 \\ 0 & 0 & 1 & 5/6 \end{array}\right] \xrightarrow{\frac{1}{5}r_2} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 21/6 \\ 0 & 1 & 0 & 11/30 \\ 0 & 0 & 1 & 5/6 \end{array}\right] \xrightarrow{r_1 - 2r_2} \boxed{\left[\begin{array}{ccc|c} 1 & 0 & 0 & 83/30 \\ 0 & 1 & 0 & 11/30 \\ 0 & 0 & 1 & 5/6 \end{array}\right] = rref[A|b].}$$

The boxed matrix corresponds to the equations $x_1 = 81/30$ and $x_2 = 11/30$ and $x_3 = 5/6$. This is the beauty of the rref, it is trivial to read off the solution from this simplified matrix.

The coefficient matrix in the above example is actually invertible and if we were to calculate A^{-1} then $Ax = b$ implies $x = A^{-1}b = (83/30, 11/30, 5/6)$. There is just one solution. The solution is unique. This is not the situation we face for the augmented coefficient matrices we need to solve for ODEs. In fact, we will expect the solution is not unique. In other words, the following example is more representative of the sort of calculation we face for systems of ODEs.

Example 4.2.3. *Consider the system of equations* $\begin{array}{c} x_1 - x_2 + x_3 = 0 \\ 3x_1 - 3x_2 = 0 \\ 2x_1 - 2x_2 - 3x_3 = 0 \end{array}$. *This gives augmented coefficient matrix* $[A|0] = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 3 & -3 & 0 & 0 \\ 2 & -2 & -3 & 0 \end{array}\right]$ *hence calculate $rref[A|0]$ as follows:*

$$[A|0] = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 3 & -3 & 0 & 0 \\ 2 & -2 & -3 & 0 \end{array}\right] \xrightarrow{r_2 - 3r_1} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & -3 & 0 \\ 2 & -2 & -3 & 0 \end{array}\right] \xrightarrow{r_3 - 2r_1}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & -5 & 0 \end{array}\right] \begin{array}{c} \xrightarrow{3r_3} \\ \xrightarrow{5r_2} \end{array} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & -15 & 0 \\ 0 & 0 & -15 & 0 \end{array}\right] \begin{array}{c} \xrightarrow{r_3 - r_2} \\ \xrightarrow{\frac{-1}{15}r_2} \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{r_1 - r_2} \boxed{\left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = rref[A|0]}$$

Note it is customary to read multiple row operations from top to bottom if more than one is listed between two of the matrices[5]. *Let's interpret the boxed reduction,*

$$x_1 - x_2 = 0, \quad x_3 = 0 \quad \Rightarrow \quad x = (x_1, x_2, x_3) = (x_2, x_2, 0) = x_2(1, 1, 0)$$

We find **infinitely many solutions**. *The set of vectors x for which $Ax = 0$ is known as the* **null space** *of A. Our calculation in this example shows $Null(A) = \{x_2(1,1,0) \mid x_2 \in \mathbb{R}\} = span\{(1,1,0)\}$.*

Definition 4.2.4. Span is set of finite linear combinations.

Given vectors $v_1, v_2, \dots v_k$ we define $span\{v_1, v_2, \dots, v_k\} = \left\{ \sum_{i=1}^{k} c_i v_i \mid c_1, \dots, c_k \in \mathbb{R} \right\}$.

[5]The multiple arrow notation should be used with caution as it has great potential to confuse. Also, you might notice that I did not strictly-speaking follow Gauss-Jordan in the operations $3r_3 \to r_3$ and $5r_2 \to r_2$. It is sometimes convenient to modify the algorithm slightly in order to avoid fractions

We can always write the solution set for a homogeneous system of linear equations as a span of appropriate vectors. Let me give another example, but I will omit the row-reduction this time. We write $A \sim B$ to indicate the matrices are **row-equivalent**. This means there is some finite set of elementary row operations which transforms A into B.

Example 4.2.5. $[A|0] = \left[\begin{array}{cccc|c} 1 & 2 & 7 & 13 & 0 \\ 3 & 6 & 7 & 11 & 0 \\ 2 & 4 & 7 & 12 & 0 \\ 5 & 10 & 7 & 9 & 0 \end{array}\right] \sim rref[A|0] = \left[\begin{array}{cccc|c} 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$. *Therefore, the system* $Ax = 0$ *with* $x = (x_1, x_2, x_3, x_4)$ *has* $x_1 + 2x_2 - x_4 = 0$ *and* $x_3 + 2x_4 = 0$ *from which we derive*

$$x = (-2x_2 + x_4, x_2, -2x_4, x_4) = x_2(-2, 1, 0, 0) + x_4(1, 0, -2, 1).$$

Thus for $A = \begin{bmatrix} 1 & 2 & 7 & 13 \\ 3 & 6 & 7 & 11 \\ 2 & 4 & 7 & 12 \\ 5 & 10 & 7 & 9 \end{bmatrix}$ *we have shown* $Null(A) = span\{(-2, 1, 0, 0), (1, 0, -2, 1)\}$.

Theorem 4.2.6.

The vectors in the null space of A are perpendicular to every row of A.

In order to describe our solution sets nicely we need to introduce the concept of **linear independence**

Definition 4.2.7. no linear dependence means the set is LI.

Given vectors $v_1, v_2, \dots v_k$ if $c_1v_1 + \cdots + c_kv_k = 0$ implies $c_1 = 0, \dots, c_k = 0$ then $\{v_1, v_2, \dots, v_k\}$ is said to be **linearly independent** or **LI**. If a set of vectors is not LI then it is **linearly dependent**.

When a set of vectors is linearly dependent it is possible to write at least one vector in the set as a linear combination of the remaining vectors in the set. Any set containing the zero vector is linearly dependent. Any set of vectors containing the same vector twice is linearly dependent. Any set of vectors containing a vector and a multiple of that vector is linearly dependent.

Definition 4.2.8. basis and dimension.

If $W = span(\beta)$ where $\beta = \{w_1, \dots, w_k\}$ is LI then we say β is a **basis** for $W$and $dim(W) = k$.

If a space is one-dimensional then it's a line, if it's two-dimensional it's a plane, if it's three-dimensional it's a volume etc. Two spaces are most interesting for linear algebra:

Definition 4.2.9. Rank & Column Space and Null Space & Nullity.

Given a matrix $A \in \mathbb{R}^{m \times n}$ we say $Null(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$ is the **null space** and **nullity** is the dimension of the null space; $dim(Null(A)) = nullity(A)$. Likewise, the **column space** of A is given by $Col(A) = span\{col_1(A), \dots, col_n(A)\}$ and we define the **rank** to be the dimension of the column space; $dim(Col(A)) = rank(A)$.

The basis for column space is given by the pivot columns whereas the basis for the null space can be derived from $rref[A|0]$ as I showed in Examples 4.2.3 and 4.2.5. Since rank is the number of pivot columns and nullity is the number of non-pivot columns (which is the number of vectors we need to form a basis for the null space) we have:

Theorem 4.2.10.

Let $A \in \mathbb{R}^{m \times n}$ then $rank(A) + nullity(A) = n$.

4.3 diagonalization and eigenvectors

We'll motivate the importance of eigenvectors later with our study of systems of differential equations. This section is just a warm-up to work on the algebra before we apply it.

Given a square matrix A we say $\vec{v} \neq 0$ is an **eigenvector** of A with **eigenvalue** λ if

$$A\vec{v} = \lambda\vec{v}$$

The matrix $I = Diag(1, 1, \dots, 1)$ is the **identity matrix** and $\vec{v} = I\vec{v}$ hence

$$A\vec{v} = \lambda I\vec{v} \quad \Rightarrow \quad (A - \lambda I)\vec{v} = 0.$$

The equation above requires $\vec{v} \in Null(A - \lambda I)$ and since $\vec{v} \neq 0$ we find $A - \lambda I$ is not an invertible matrix. This means that the determinant of $A - \lambda I$ must be zero. We have a name for this equation:

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

is the **characteristic equation** of A and $p_A(t) = det(A - tI)$ is the **characteristic polynomial** of A. So, the eigenvalues of A are the roots of the characteristic equation. This is an n-th order polynomial equation so there are always n-solutions, possibly repeated, possibly complex, by the Fundamental Theorem of Algebra. To find the eigenvalues of A we need to solve the characteristic equation.

Definition 4.3.1. *eigenvalues and eigenvectors*

Suppose A is a real $n \times n$ matrix then we say $\lambda \in \mathbb{R}$ which is a solution of $det(A - \lambda I) = 0$ is an **eigenvalue of** A. Given such an eigenvalue λ a nonzero real vector $\vec{v}$ such that $(A - \lambda I)\vec{v} = 0$ is called an **eigenvector** with eigenvalue λ.

Likewise, $\lambda \in \mathbb{C}$ which is a solution of $det(A - \lambda I) = 0$ is an **complex eigenvalue** of A and the complex vector $\vec{v} \neq 0$ for which $(A - \lambda I)\vec{v} = 0$ is a **complex eigenvector** of A.

Example 4.3.2. *Let* $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ *then calculate:*

$$det(A - \lambda I) = det \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} = (\lambda - 1)^2 - 4 = (\lambda + 1)(\lambda - 3).$$

We deduce eigenvalues are $\lambda_1 = -1$ *and* $\lambda_2 = 3$.

$$(A + I)\vec{v} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\vec{v} = 0 \quad \Rightarrow \quad \vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Thus $\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *is an eigenvector with eigenvalue* $\lambda_1 = -1$. *Next, study* $\lambda_2 = 3$,

$$(A - 3I)\vec{v} = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}\vec{v} = 0 \quad \Rightarrow \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Hence $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *is an eigenvector with eigenvalue* $\lambda_2 = 3$.

Let us continue our discussion of $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. We found a pair of representative eigenvectors, but these solutions are by no means unique. I simply chose them since I like to use integer components if possible. In fact, $\begin{bmatrix} c \\ -c \end{bmatrix} \in Null(A+I)$ for any scalar c. We call $Null(A+I)$ the $\lambda_1 = -1$ **eigenspace** of A. Finding an eigenvector is the calculation of finding the null space basis, however, with these 2×2 problems we can usually just guess the solution of $(A-\lambda I)\vec{u}=0$ by inspection of the matrix. We just pick a vector perpendicular to both rows of $A-\lambda I$. Note, every vector in the $\lambda_2 = 3$ eigenspace is a multiple of $\vec{v}_2$ hence $Null(A-3I) = span\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$. Generally speaking, we need to learn how to find the basis for each eigenspace of A in order to solve systems of differential equations.

I'll continue to study our nice example to illustrate the problem of diagonalization of a matrix. Let me first share the theory involved:

Theorem 4.3.3.

Let $A \in \mathbb{R}^{n\times n}$ with linearly independent eigenvectors $v_1, v_2, \dots, v_n$ with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ possibly repeated then if $P = [v_1|v_2|\cdots|v_n]$ we have

$$P^{-1}AP = Diag(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Notice we are not generally able to find a set of n-LI eigenvectors for a given matrix. For example, no vector satisfies $Rv = \lambda v$ if you think about R as the rotation matrix in the plane of some angle other than $0, \pi$ radians. That said, one special class of matrices allows us to find really nice eigenvectors with all real eigenvalues.

Theorem 4.3.4. *Real Spectral Theorem:*

Let $A \in \mathbb{R}^{n\times n}$. If $A^T = A$ then there exists an orthonormal (meaning $u_i \bullet u_j = \delta_{ij}$, they are perpendicular to one another and all of length one) set of eigenvectors $u_1, u_2, \dots, u_n$ whose eigenvalues are real and for which $Q = [u_1|u_2|\dots|u_n]$ is a matrix satisfying $Q^TQ = I$ hence $Q^T = Q^{-1}$ and

$$Q^TAQ = Diag(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Example 4.3.5. *Notice* $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ *has* $A^T = A$. *Normalizing* $\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *gives orthonormal eigenvectors* $\vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *and* $\vec{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *then we construct*

$$Q = [\vec{u}_1|\vec{u}_2] = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

and

$$Q^TAQ = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 3 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$$

The process of transforming a matrix to a corresponding diagonal matrix as shown in the above example is known as **diagonalization**. This is closely tied to the problem of **uncoupling expressions** which requires a change of variables to remove **cross-terms**. Let's see how this works for a quadratic form which is based on our example matrix.

Example 4.3.6. *Let* $Q(x,y) = [x,y]^T \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 4xy + y^2$. *If we were to ask most students of highschool algebra to graph* $x^2+4xy+y^2 = 1$ *then this would probably be a tough question. Let's make a substitution based on the orthonormal eigenbasis we found in Example 4.3.5. Define* **eigencoordinates** $\bar{x}, \bar{y}$ *by*

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = Q^T \begin{bmatrix} x \\ y \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} x \\ y \end{bmatrix} = Q \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \quad \& \quad [\bar{x}, \bar{y}]Q^T = [x,y].$$

Therefore, we can change coordinates for the quadratic form via:

$$Q(x,y) = [x,y]A \begin{bmatrix} x \\ y \end{bmatrix} = [\bar{x},\bar{y}]Q^T AQ \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = [\bar{x},\bar{y}] \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = -\bar{x}^2 + 3\bar{y}^2.$$

Now we can easily understand $x^2 + 4xy + y^2 = 1$ *is just* $-\bar{x}^2 + 3\bar{y}^2 = 1$. *The eigencoordinates have revealed the mystery curve is a hyperbola which opens vertically with respect to the* $\bar{x}, \bar{y}$*-coordinate system*[6]

I include the example above since I want you to appreciate that eigenvectors help us do much more than just solve systems of differential equations. I'll give one more example before we get back to the problem of differential equations.

Example 4.3.7. *Let* $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

$$\begin{aligned} det(A - \lambda I) &= det \begin{bmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{bmatrix} \\ &= (1-\lambda)[(1-\lambda)^2 - 1] - [(1-\lambda) - 1] + [1 - (1-\lambda)] \\ &= -\lambda^2(\lambda - 3). \end{aligned}$$

We find eigenvalues $\lambda_1 = 0$ *and* $\lambda_2 = 3$. *Consider:*

$$[A - 3I|0] = \left[\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array}\right] \xrightarrow{r_1 + 2r_2,\ r_3 - r_2} \left[\begin{array}{ccc|c} 0 & -3 & 3 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 3 & -3 & 0 \end{array}\right]$$

$$\xrightarrow{r_3 + r_1} \left[\begin{array}{ccc|c} 0 & -3 & 3 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{-r_1/3} \left[\begin{array}{ccc|c} 0 & 1 & -1 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{r_2 + 2r_1} \left[\begin{array}{ccc|c} 0 & 1 & -1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $(A - 3I)\vec{u} = 0$ *with* $\vec{u} = (u_1, u_2, u_3)$ *must have* $u_1 = u_3$ *and* $u_2 = u_3$ *which gives* $\vec{u} = (u_3, u_3, u_3) = u_3(1,1,1)$. *In short, a nice eigenvector with eigenvalue* 3 *is the vector* $\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. *I would add, we certainly could have guess this without all this calculation. The zero eigenvalue*

[6] if you want to understand more, think about how $x = 0$ and $y = 0$ define the y-axis and x-axis respectively. Thus, $\bar{x} = 0$ gives the $\bar{y}$-axis and $\bar{y} = 0$ gives the $\bar{x}$-axis.

requires less calculation with row reduction, but the logic is a bit trickier. We seek to find solutions of $A\vec{u} = 0$ where $\vec{u} \neq 0$. Consider,

$$[A|0] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right] \xrightarrow{r_2 - r_1,\ r_3 - r_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

If we let $\vec{u} = (u_1, u_2, u_3)$ then we see that $A\vec{u} = 0$ requires that $u_1 + u_2 + u_3 = 0$ hence $u_1 = -u_2 - u_3$ (it is traditional and customary to solve for the so-called **pivotal variable** *as functions of the* **non-pivotal variables** *). Hence,*

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -u_2 - u_3 \\ u_2 \\ u_3 \end{bmatrix} = u_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + u_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

We find the eigenvalue $\lambda_2 = 0$ has a pair of linearly independent eigenvectors $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.
In summary, we found the $\lambda_1 = 3$ eigenspace $Null(A - 3I) = span\{(1,1,1)\}$ whereas the $\lambda_2 = 0$ eigenspace has $Null(A) = span\{(-1,1,0), (-1,0,1)\}$. If we define $v_1 = (1,1,1)$ and $v_2 = (-1,1,0)$ and $v_3 = (-1,0,1)$. Then $\beta = \{v_1, v_2, v_3\}$ is an eigenbasis for A and we can calculate,

$$[\beta] = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \Rightarrow \quad [\beta]^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

I used technology to calculate the inverse above. Then the diagonalization of A is seen as follows:

$$[\beta]^{-1} A [\beta] = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

To complete the thought, if we define $(\bar{x}, \bar{y}, \bar{z}) = [\beta]^{-1}(x, y, z)$ then we would find

$$\begin{aligned} x^2 + y^2 + z^2 + 2xy + 2xz + 2yz &= [x, y, z]^T \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= [x, y, z]^T A(x, y, z) \\ &= [x, y, z]^T [\beta][\beta]^{-1} A [\beta][\beta]^{-1}(x, y, z) \\ &= [\bar{x}, \bar{y}, \bar{z}] \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} (\bar{x}, \bar{y}, \bar{z}) \\ &= 3\bar{x}^2. \end{aligned}$$

This illustrates what I mean by **decoupling** *the original formula in x, y, z involves mixed terms like $2xy, 2xz, 2yz$ whereas the final eigenformula is much simplified with no coupled terms.*

4.4 the normal form and theory for systems

A system of ODEs in normal form is a finite collection of first order ODEs which share dependent variables and a single independent variable.

1. $(n = 1)$ $\frac{dx}{dt} = A_{11}x + f$

2. $(n = 2)$ $\frac{dx}{dt} = A_{11}x + A_{12}y + f_1$ and $\frac{dy}{dt} = A_{21}x + A_{22}y + f_2$ we can express this in **matrix normal form** as follows, use $x = x_1$ and $y = x_2$,

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

   This is nicely abbreviated by writing $d\vec{x}/dt = A\vec{x} + \vec{f}$ where $\vec{x} = (x_1, x_2)$ and $\vec{f} = (f_1, f_2)$ whereas the 2×2 matrix A is called the **coefficient matrix** of ths system.

3. $(n = 3)$ The matrix normal form is simply

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

   Expanded into **scalar normal form** we have $\frac{dx_1}{dt} = A_{11}x_1 + A_{12}x_2 + A_{13}x_3 + f_1$ and $\frac{dx_2}{dt} = A_{21}x_1 + A_{22}x_2 + A_{23}x_3 + f_2$ and $\frac{dx_3}{dt} = A_{31}x_1 + A_{32}x_2 + A_{33}x_3 + f_3$.

Generally an n-th order system of ODEs in normal form on an interval $I \subseteq \mathbb{R}$ can be written as $\frac{dx_i}{dt} = \sum_{j=1}^{n} A_{ij}x_j + f_i$ for **coefficient functions** $A_{ij} : I \subseteq \mathbb{R} \to \mathbb{R}$ and **forcing functions** $f_i : I \subseteq \mathbb{R} \to \mathbb{R}$. You might consider the problem of solving a system of k-first order differential equations in n-dependent variables where $n \neq k$, however, we do not discuss such over or underdetermined problems in these notes. That said, the concept of a system of differential equations in normal form is perhaps more general than you expect. Let me illustrate this by example. I'll start with a single second order ODE:

Example 4.4.1. *Consider $ay'' + by' + cy = f$. We define $x_1 = y$ and $x_2 = y'$. Observe that*

$$x_1' = x_2 \qquad \& \qquad x_2' = y'' = \frac{1}{a}(f - by' - cy) = \frac{1}{a}(f - bx_2 - cx_1)$$

Thus,

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -c/a & -b/a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ f/a \end{bmatrix}$$

The matrix $\begin{bmatrix} 0 & 1 \\ -c/a & -b/a \end{bmatrix}$ is called the **companion matrix** *of the second order ODE $ay'' + by' + cy = f$.*

The example above nicely generalizes to the general n-th order linear ODE.

Example 4.4.2. *Consider $a_o y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = f$. Introduce variables to reduce the order:*

$$x_1 = y, \qquad x_2 = y', \qquad x_3 = y'', \qquad \dots \qquad x_n = y^{(n-1)}$$

From which is is clear that $x_1' = x_2$ and $x_2' = x_3$ continuing up to $x_{n-1}' = x_n$ and $x_n' = y^{(n)}$. Hence,

$$x_n' = -\frac{a_1}{a_o}x_n - \cdots - \frac{a_{n-1}}{a_o}x_2 - \frac{a_n}{a_o}x_1 + f$$

Once again the matrix below is called the **companion matrix** *of the given n-th order ODE.*

$$\begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_{n-1}' \\ x_n' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\frac{a_n}{a_o} & -\frac{a_{n-1}}{a_o} & -\frac{a_{n-2}}{a_o} & \cdots & -\frac{a_2}{a_o} & -\frac{a_1}{a_o} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{f}{a_o} \end{bmatrix}$$

The problem of many higher order ODEs is likewise confronted by introducing variables to reduce the order.

Example 4.4.3. *Consider $y'' + 3x' = \sin(t)$ and $x'' + 6y' - x = e^t$. We begin with a system of two second order differential equations. Introduce new variables:*

$$x_1 = x, \quad x_2 = y, \quad x_3 = x', \quad x_4 = y'$$

It follows that $x_3' = x''$ and $x_4' = y''$ whereas $x_1' = x_3$ and $x_2' = x_4$. We convert the given differential equations to first order ODEs:

$$x_4' + 3x_3 = \sin(t) \quad \& \quad x_3' + 6x_4 - x_1 = e^t$$

Let us collect these results as a matrix problem:

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -6 \\ 0 & 0 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ e^t \\ \sin(t) \end{bmatrix}$$

Generally speaking the order of the normal form corresponding to a system of higher order ODE will simply be the sum of the orders of the systems (assuming the given system has no reundancies; for example $x'' + y'' = x$ and $x'' - x = -y''$ are redundant). I will not prove the following assertion, however, it should be fairly clear why it is true given the examples thus far discussed:

Proposition 4.4.4. *linear systems have a normal form.*

A given systems of linear ODEs may be converted to an equivalent system of first order ODEs in normal form.

For this reason the first order problem will occupy the majority of our time. That said, the method of the next section is applicable to any order.

Since normal forms are essentially general it is worthwhile to state the theory which will guide our work. I do not offer all the proof here, but you can find proof in many texts. For example, in Nagel Saff and Snider these theorems are given in §9.4 and are proved in Chapter 13.

Definition 4.4.5. *linear independence of vector-valued functions*

Suppose $\vec{v}_j : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ is a function for $j = 1, 2, \dots, k$ then we say that $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ is linearly independent on I iff $\sum_{j=1}^{k} c_j \vec{v}_j(t) = 0$ for all $t \in I$ implies $c_j = 0$ for $j = 1, 2, \dots, k$.

We can use the determinant to test LI of a set of n-vectors which are all n-dimensional vectors. It is true that $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is LI on I iff $det[\vec{v}_1(t)|\vec{v}_2(t)| \cdots |\vec{v}_n(t)| \neq 0$ for all $t \in I$.

Definition 4.4.6. *wronskian for vector-valued functions of a real variable.*

Suppose $\vec{v}_j : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ is differentiable for $j = 1, 2, \dots, n$. The **Wronskian** is defined by $W(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n; t) = det[\vec{v}_1|\vec{v}_2| \dots |\vec{v}_n]$ for each $t \in I$.

Theorems for wronskians of solutions sets mirror those already discussed for the n-th order problem.

Definition 4.4.7. *solution and homogeneous solutions of $d\vec{x}/dt = A\vec{x} + \vec{f}$*

Let $A : I \rightarrow \mathbb{R}^{n \times n}$ and $\vec{f} : I \rightarrow \mathbb{R}^n$ be continuous. A **solution of** $d\vec{v}/dt = A\vec{v} + \vec{f}$ **on** $I \subseteq \mathbb{R}$ is a vector-valued function $\vec{x} : I \rightarrow \mathbb{R}^n$ such that $d\vec{x}/dt = A\vec{x} + \vec{f}$ for all $t \in I$. A **homogeneous solution** on $I \subseteq \mathbb{R}$ is a solution of $d\vec{v}/dt = A\vec{v}$.

In the example below we see three LI homogeneous solutions and a single particular solution.

Example 4.4.8. *Suppose $x' = x - 1$, $y' = 2y - 2$ and $z' = 3z - 3$. In matrix normal form we face:*

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}$$

It is easy to show by separately solving the the DEqns that $x = c_1 e^t + 1$, $y = c_2 e^{2t} + 1$ and $z = c_3 e^{3t} + 1$. In vector notation the solution is

$$\vec{x}(t) = \begin{bmatrix} c_1 e^t + 1 \\ c_2 e^{2t} + 1 \\ c_3 e^{3t} + 1 \end{bmatrix} = c_1 \begin{bmatrix} e^t \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ e^{2t} \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ e^{3t} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

I invite the reader to show that $S = \{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ is LI on $\mathbb{R}$ where $\vec{x}_1(t) = \langle e^t, 0, 0 \rangle$, $\vec{x}_2(t) = \langle 0, e^{2t}, 0 \rangle$ and $\vec{x}_3(t) = \langle 0, 0, e^{3t} \rangle$. On the other hand, $\vec{x}_p = \langle 1, 1, 1 \rangle$ is a particular solution to the given problem.

In truth, any choice of c_1, c_2, c_3 with at least one nonzero constant will produce a homogeneous solution. To obtain the solutions I pointed out in the example you can choose $c_1 = 1, c_2 = 0, c_3 = 0$ to obtain $\vec{x}_1(t) = \langle e^t, 0, 0 \rangle$ or $c_1 = 0, c_2 = 1, c_3 = 0$ to obtain $\vec{x}_2(t) = \langle 0, e^{2t}, 0 \rangle$ or $c_1 = 0, c_2 = 0, c_3 = 1$ to obtain $\vec{x}_3(t) = \langle 0, 0, e^{3t} \rangle$.

Definition 4.4.9. *fundamental solution set of a linear system $d\vec{x}/dt = A\vec{x} + \vec{f}$*

Let $A : I \rightarrow \mathbb{R}^{n \times n}$ and $\vec{f} : I \rightarrow \mathbb{R}^n$ be continuous. A **fundmental solution set** on $I \subseteq \mathbb{R}$ is a set of n-homogeneous solutions of $d\vec{v}/dt = A\vec{v} + \vec{f}$ for which $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ is a LI set on I. A **solution matrix** on $I \subseteq \mathbb{R}$ is a matrix X is a matrix for which each column is a homogeneous solution on I. A **fundamental matrix** on $I \subseteq \mathbb{R}$ is an invertible solution matrix.

Example 4.4.10. *Continue Example 4.4.8. Note that $S = \{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ is a fundamental solution set. The fundamental solution matrix is found by concatenating $\vec{x}_1, \vec{x}_2$ and $\vec{x}_3$:*

$$X = [\vec{x}_1|\vec{x}_2|\vec{x}_3] = \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{3t} \end{bmatrix}$$

Observe $det(X) = e^t e^{2t} e^{3t} = e^{6t} \neq 0$ on $\mathbb{R}$ hence X is invertible on $\mathbb{R}$.

Example 4.4.11. *Let $A = \begin{bmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix}$ define the system of DEqns $\frac{d\vec{x}}{dt} = A\vec{x}$. I claim that the matrix $X(t) = \begin{bmatrix} 0 & -e^{4t} & -2e^{2t} \\ e^{4t} & 0 & e^{2t} \\ 0 & e^{4t} & e^{2t} \end{bmatrix}$ is a solution matrix. Calculate,*

$$AX = \begin{bmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix} \begin{bmatrix} 0 & -e^{4t} & -2e^{2t} \\ e^{4t} & 0 & e^{2t} \\ 0 & e^{4t} & e^{2t} \end{bmatrix} = \begin{bmatrix} 0 & -4e^{4t} & -4e^{2t} \\ 4e^{4t} & 0 & 2e^{2t} \\ 0 & 4e^{4t} & 2e^{2t} \end{bmatrix}.$$

On the other hand, differentiation yields $X' = \begin{bmatrix} 0 & -4e^{4t} & -4e^{2t} \\ 4e^{4t} & 0 & 2e^{2t} \\ 0 & 4e^{4t} & 2e^{2t} \end{bmatrix}$. Therefore $X' = AX$. Notice that if we express X in terms of its columns $X = [\vec{x}_1|\vec{x}_2|\vec{x}_3]$ then it follows that $X' = [\vec{x}_1{}'|\vec{x}_2{}'|\vec{x}_3{}']$ and $AX = A[\vec{x}_1|\vec{x}_2|\vec{x}_3] = [A\vec{x}_1|A\vec{x}_2|A\vec{x}_3]$ hence

$$\vec{x}_1{}' = A\vec{x}_1 \quad \& \quad \vec{x}_2{}' = A\vec{x}_2 \quad \& \quad \vec{x}_3{}' = A\vec{x}_3$$

We find that $\vec{x}_1(t) = \langle 0, e^{4t}, 0\rangle$, $\vec{x}_2(t) = \langle -e^{4t}, 0, e^{4t}\rangle$ and $\vec{x}_3(t) = \langle -2e^{2t}, e^{2t}, e^{2t}\rangle$ form a fundamental solution set for the given system of DEqns.

Theorem 4.4.12. *Let $A : I \to \mathbb{R}^{n\times n}$ and $\vec{f} : I \to \mathbb{R}^n$ be continuous.*

1. there exists a fundamental solution set $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ on I
2. if $t_o \in I$ and $\vec{x}_o$ is a given initial condition vector then there exists a unique solution $\vec{x}$ on I such that $\vec{x}(t_o) = \vec{x}_o$
3. the **general solution** has the form $\vec{x} = \vec{x}_h + \vec{x}_p$ where $\vec{x}_p$ is a **particular solution** and $\vec{x}_h$ is the **homogeneous solution** is formed by a real linear combination of the fundamental solution set ($\vec{x}_h = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n$)

The term *general solution* is intended to indicate that the formula given includes all possible solutions to the problem. Part (2.) of the theorem indicates that there must be some 1-1 correspondance between a given initial condition and the choice of the constants $c_1, c_2, \dots, c_n$ with respect to a given fundamental solution set. Observe that if we define $\vec{c} = [c_1, c_2, \dots, c_n]^T$ and the fundamental matrix $X = [\vec{x}_1|\vec{x}_2|\cdots|\vec{x}_n]$ we can express the homogeneous solution via a matrix-vector product:

$$\vec{x}_h = X\vec{c} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n \qquad \Rightarrow \qquad \boxed{\vec{x}(t) = X(t)\vec{c} + \vec{x}_p(t)}$$

Further suppose that we wish to set $\vec{x}(t_o) = \vec{x}_o$. We need to solve for $\vec{c}$:

$$\vec{x}_o = X(t_o)\vec{c} + \vec{x}_p(t_o) \quad \Rightarrow \quad X(t_o)\vec{c} = \vec{x}_o - \vec{x}_p(t_o)$$

Since $X^{-1}(t_o)$ exists we can multiply by the inverse on the right and find

$$\vec{c} = X^{-1}(t_o)\big[\vec{x}_o - \vec{x}_p(t_o)\big]$$

Next, place the result above back in the general solution to derive

$$\boxed{\vec{x}(t) = X(t)X^{-1}(t_o)\big[\vec{x}_o - \vec{x}_p(t_o)\big] + \vec{x}_p(t)}$$

We can further simplify this general formula in the constant coefficient case, or in the study of variation of parameters for systems. Note that in the homogeneous case this gives us a clean formula to calculate the constants to fit initial data:

$$\boxed{\vec{x}(t) = X(t)X^{-1}(t_o)\vec{x}_o} \qquad \text{(homogeneous case)}$$

Example 4.4.13. *We found $x' = -y$ and $y' = x$ had solutions $x(t) = c_1\cos(t) + c_2\sin(t)$ and $y(t) = c_1\sin(t) - c_2\cos(t)$. It follows that $X(t) = \begin{bmatrix} \cos(t) & \sin(t) \\ \sin(t) & -\cos(t) \end{bmatrix}$. Calculate that $det(X) = -1$ to see that $X^{-1}(t) = \begin{bmatrix} \cos(t) & \sin(t) \\ \sin(t) & -\cos(t) \end{bmatrix}$. Suppose we want the solution through (a, b) at time t_o then the solution is given by*

$$\vec{x}(t) = X(t)X^{-1}(t_o)\vec{x}_o = \begin{bmatrix} \cos(t) & \sin(t) \\ \sin(t) & -\cos(t) \end{bmatrix}\begin{bmatrix} \cos(t_o) & \sin(t_o) \\ \sin(t_o) & -\cos(t_o) \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}.$$

This concludes our brief tour of the theory for systems of ODEs. Clearly we have two main goals past this point (1.) find the fundamental solution set (2.) find the particular solution.

4.5 solutions by eigenvector

We narrow our focus at this point: our goal is to find nontrivial[7] solutions to the homogeneous constant coefficient problem $\frac{d\vec{x}}{dt} = A\vec{x}$ where $A \in \mathbb{R}^{\,n\times n}$. A reasonable ansatz for this problem is that the solution should have the form $\vec{x} = e^{\lambda t}\vec{u}$ for some constant scalar λ and some constant vector $\vec{u}$. If such solutions exist then what conditions must we place on λ and $\vec{u}$? To begin clearly $\vec{u} \neq 0$ since we are seeking nontrivial solutions. Differentiate,

$$\frac{d}{dt}\big[e^{\lambda t}\vec{u}\big] = \frac{d}{dt}\big[e^{\lambda t}\big]\vec{u} = \lambda e^{\lambda t}\vec{u}$$

Hence $\frac{d\vec{x}}{dt} = A\vec{x}$ implies $\lambda e^{\lambda t}\vec{u} = Ae^{\lambda t}\vec{u}$. However, $e^{\lambda t} \neq 0$ hence we find $\lambda\vec{u} = A\vec{u}$. We can write the vector $\lambda\vec{u}$ as a matrix product with identity matrix I; $\lambda\vec{u} = \lambda I\vec{u}$. Therefore, we find

$$\boxed{(A - \lambda I)\vec{u} = 0}$$

[7]nontrivial simply means the solution is not identically zero. The zero solution does exist, but it is not the solution we are looking for...

to be a necessary condition for the solution. Note that the system of linear equations defined by $(A-\lambda I)\vec{u}=0$ is consistent since 0 is a solution. It follows that for $\vec{u}\neq 0$ to be a solution we must have that the matrix $(A-\lambda I)$ is singular. It follows that we find

$$\boxed{det(A-\lambda I)=0}$$

a necessary condition for our solution. Moreover, for a given matrix A this is nothing more than an n-th order polynomial in λ hence there are at most n-distinct solutions for λ. The equation $det(A-\lambda I)=0$ is called the **characteristic equation** of A and the solutions are called **eigenvalues**. The nontrivial vector $\vec{u}$ such that $(A-\lambda I)\vec{u}=0$ is called the **eigenvector** with **eigenvalue** λ. We often abbreviate these by referring to "e-vectors" or "e-values". Many interesting theorems are known for eigenvectors, see a linear algebra text or my linear notes for elaboration on this point.

Example 4.5.1. Problem: *find the fundamental solutions of the system* $x'=-4x-y$ *and* $y'=5x+2y$

Solution: *we seek to solve* $\frac{d\vec{x}}{dt}=A\vec{x}$ *where* $A=\begin{bmatrix}-4 & -1\\ 5 & 2\end{bmatrix}$. *Consider the characteristic equation:*

$$\begin{aligned} det(A-\lambda I) &= det\begin{bmatrix}-4-\lambda & -1\\ 5 & 2-\lambda\end{bmatrix} \\ &= (-4-\lambda)(2-\lambda)+5 \\ &= \lambda^2+2\lambda-3 \\ &= (\lambda+3)(\lambda-1) \\ &= 0 \end{aligned}$$

We find $\lambda_1=1$ *and* $\lambda_2=-3$. *Next calculate the e-vectors for each e-value. We seek* $\vec{u}_1=[u,v]^T$ *such that* $(A-I)\vec{u}_1=0$ *thus solve:*

$$\begin{bmatrix}-5 & -1\\ 5 & 1\end{bmatrix}\begin{bmatrix}u\\ v\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix} \quad\Rightarrow\quad -5u-v=0 \quad\Rightarrow\quad v=-5u,\ u\neq 0 \quad\Rightarrow\quad \vec{u}_1=\begin{bmatrix}u\\ -5u\end{bmatrix}$$

Naturally we can write $\vec{u}_1=u[1,-5]^T$ *and for convenience we set* $u=1$ *and find* $\vec{u}_1=[1,-5]^T$ *which gives us the fundamental solution* $\boxed{\vec{x}_1(t)=e^t[1,-5]^T}$. *Continue*[8] *to the next e-value* $\lambda_2=-3$ *we seek* $\vec{u}_2=[u,v]^T$ *such that* $(A+3I)\vec{u}_2=0$.

$$\begin{bmatrix}-1 & -1\\ 5 & 5\end{bmatrix}\begin{bmatrix}u\\ v\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix} \quad\Rightarrow\quad -u-v=0 \quad\Rightarrow\quad v=-u,\ u\neq 0 \quad\Rightarrow\quad \vec{u}_2=\begin{bmatrix}u\\ -u\end{bmatrix}$$

Naturally we can write $\vec{u}_2=u[1,-1]^T$ *and for convenience we set* $u=1$ *and find* $\vec{u}_2=[1,-1]^T$ *which gives us the fundamental solution* $\boxed{\vec{x}_2(t)=e^{-3t}[1,-1]^T}$. *The fundmental solution set is given by* $\{\vec{x}_1,\vec{x}_2\}$ *and the domains of these solution clearly extend to all of* $\mathbb{R}$.

[8] the upcoming u,v are not the same as those I just worked out, I call these letters disposable variables because I like to reuse them in several ways in a particular example where we repeat the e-vector calculation over several e-values.

We can assemble the general solution as a linear combination of the fundamental solutions $\vec{x}(t) = c_1\vec{x}_1 + c_2\vec{x}_2$. In particular this yields

$$\vec{x}(t) = c_1\vec{x}_1 + c_2\vec{x}_2 = c_1e^t\begin{bmatrix} 1 \\ -5 \end{bmatrix} + c_2e^{-3t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} c_1e^t + c_2e^{-3t} \\ -5c_1e^t - c_2e^{-3t} \end{bmatrix}$$

Thus the system $x' = -4x - y$ and $y' = 5x + 2y$ has **scalar** solutions $x(t) = c_1e^t + c_2e^{-3t}$ and $y(t) = -5c_1e^t - c_2e^{-3t}$. Finally, a fundamental matrix for this problem is given by

$$X = [\vec{x}_1|\vec{x}_2] = \begin{bmatrix} e^t & e^{-3t} \\ -5e^t & -e^{-3t} \end{bmatrix}.$$

Example 4.5.2. Problem: *find the fundamental solutions of the system $x' = -3x$ and $y' = -3y$*

Solution: *we seek to solve $\frac{d\vec{x}}{dt} = A\vec{x}$ where $A = \begin{bmatrix} -3 & 0 \\ 0 & -3 \end{bmatrix}$. Consider the characteristic equation:*

$$det(A - \lambda I) = det\begin{bmatrix} -3-\lambda & 0 \\ 0 & -3-\lambda \end{bmatrix} = (\lambda+3)^2 = 0$$

We find $\lambda_1 = -3$ and $\lambda_2 = -3$. Finding the eigenvectors here offers an unusual algebra problem; to find $\vec{u}$ with e-value $\lambda = -3$ we should find nontrivial solutions of $(A+3I)\vec{u} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = 0$. We find no condition on $\vec{u}$. It follows that **any** *nonzero vector is an eigenvector of A. Indeed, note that $A = -3I$ and $A\vec{u} = -3I\vec{u}$ hence $(A+3I)\vec{u} = 0$. Convenient choices for $\vec{u}$ are $[1,0]^T$ and $[0,1]^T$ hence we find fundamental solutions:*

$$\vec{x}_1(t) = e^{-3t}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} e^{-3t} \\ 0 \end{bmatrix} \qquad \& \qquad \vec{x}_2(t) = e^{-3t}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ e^{-3t} \end{bmatrix}.$$

We can assemble the general solution as a linear combination of the fundamental solutions $\vec{x}(t) = c_1\begin{bmatrix} e^{-3t} \\ 0 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ e^{-3t} \end{bmatrix}$. Thus the system $x' = -3x$ and $y' = -3y$ has **scalar** solutions $x(t) = c_1e^{-3t}$ and $y(t) = c_2e^{-3t}$. Finally, a fundamental matrix for this problem is given by

$$X = [\vec{x}_1|\vec{x}_2] = \begin{bmatrix} e^{-3t} & 0 \\ 0 & e^{-3t} \end{bmatrix}.$$

Example 4.5.3. Problem: *find the fundamental solutions of the system $x' = 3x + y$ and $y' = -4x - y$*

Solution: *we seek to solve $\frac{d\vec{x}}{dt} = A\vec{x}$ where $A = \begin{bmatrix} 3 & 1 \\ -4 & -1 \end{bmatrix}$. Consider the characteristic equation:*

$$\begin{aligned} det(A - \lambda I) &= det\begin{bmatrix} 3-\lambda & 1 \\ -4 & -1-\lambda \end{bmatrix} \\ &= (\lambda-3)(\lambda+1) + 4 \\ &= \lambda^2 - 2\lambda + 1 \\ &= (\lambda-1)^2 \\ &= 0 \end{aligned}$$

We find $\lambda_1 = 1$ *and* $\lambda_2 = 1$. *Let us find the e-vector* $\vec{u}_1 = [u, v]^T$ *such that* $(A - I)\vec{u}_1 = 0$

$$\begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad 2u + v = 0 \quad \Rightarrow \quad v = -2u, \ u \neq 0 \quad \Rightarrow \quad \vec{u}_1 = \begin{bmatrix} u \\ -2u \end{bmatrix}$$

We choose $u = 1$ *for convenience and thus find the fundamental solution* $\vec{x}_1(t) = e^t \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

Remark 4.5.4.

In the previous example the **algebraic multiplicity** of the e-value $\lambda = 1$ was 2. However, we found only one LI e-vector. This means the **geometric multiplicity** for $\lambda = 1$ is only 1. This means we are missing a vector and hence a fundamental solution. Where is $\vec{x}_2$ which is LI from the $\vec{x}_1$ we just found? This question is ultimately answered via the matrix exponential.

Example 4.5.5. Problem: *find the fundamental solutions of the system* $x' = -y$ *and* $y' = 4x$

Solution: *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix}$. *Consider the characteristic equation:*

$$det(A - \lambda I) = det \begin{bmatrix} -\lambda & -1 \\ 4 & -\lambda \end{bmatrix} = \lambda^2 + 4 = 0 \quad \Rightarrow \quad \lambda = \pm 2i.$$

This e-value is a **pure imaginary** *number which is a special type of* **complex number** *where there is no real part. Careful review of the arguments that framed the e-vector solution reveal that the same calculations apply when either* λ *or* $\vec{u}$ *are complex. With this in mind we seek the e-vector for* $\lambda = 2i$: *let us find the e-vector* $\vec{u}_1 = [u, v]^T$ *such that* $(A - 2iI)\vec{u}_1 = 0$

$$\begin{bmatrix} -2i & -1 \\ 4 & -2i \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad -2iu - v = 0 \quad \Rightarrow \quad v = -2iu, \ u \neq 0 \quad \Rightarrow \quad \vec{u}_1 = \begin{bmatrix} u \\ -2iu \end{bmatrix}$$

Let $u = 1$ *for convenience and find* $\vec{u}_1 = [1, -2i]^T$. *We find the* **fundamental complex solution** $\vec{x}$:

$$\vec{x} = e^{2it} \begin{bmatrix} 1 \\ -2i \end{bmatrix} = \big(\cos(2t) + i\sin(2t)\big) \begin{bmatrix} 1 \\ -2i \end{bmatrix} = \begin{bmatrix} \cos(2t) + i\sin(2t) \\ -2i\cos(2t) + 2\sin(2t) \end{bmatrix}$$

Note: if $\vec{x} = Re(\vec{x}) + iIm(\vec{x})$ *then it follows that the real and imaginary parts of the complex solution are themselves real solutions. Why? Because differentiation with respect to* t *is defined such that:*

$$\frac{d\vec{x}}{dt} = \frac{dRe(\vec{x})}{dt} + i\frac{dIm(\vec{x})}{dt}$$

and $A\vec{x} = A[Re(\vec{x}) + iIm(\vec{x})] = A\,Re(\vec{x}) + iA\,Im(\vec{x})$. *However, we know* $d\vec{x}/dt = A\vec{x}$ *hence we find, equating real parts and imaginary parts separately that:*

$$\frac{dRe(\vec{x})}{dt} = ARe(\vec{x}) \qquad \& \qquad \frac{dIm(\vec{x})}{dt} = AIm(\vec{x})$$

Hence $\vec{x}_1 = Re(\vec{x})$ *and* $\vec{x}_2 = Im(\vec{x})$ *give a solution set for the given system. In particular we find the fundamental solution set*

$$\vec{x}_1(t) = \begin{bmatrix} \cos(2t) \\ 2\sin(2t) \end{bmatrix} \qquad \& \qquad \vec{x}_2(t) = \begin{bmatrix} \sin(2t) \\ -2\cos(2t) \end{bmatrix}.$$

We can assemble the general solution as a linear combination of the fundamental solutions $\vec{x}(t) = c_1 \begin{bmatrix} \cos(2t) \\ 2\sin(2t) \end{bmatrix} + c_2 \begin{bmatrix} \sin(2t) \\ -2\cos(2t) \end{bmatrix}$. Thus the system $x' = -y$ and $y' = 4x$ has **scalar** solutions
$x(t) = c_1 \cos(2t) + c_2 \sin(2t)$ and $y(t) = 2c_1 \sin(2t) - 2c_2 \cos(2t)$. Finally, a fundamental matrix for this problem is given by

$$X = [\vec{x}_1|\vec{x}_2] = \begin{bmatrix} \cos(2t) & \sin(2t) \\ 2\sin(2t) & -2\cos(2t) \end{bmatrix}.$$

Example 4.5.6. Problem: *find the fundamental solutions of the system* $x' = 2x - y$ *and* $y' = 9x + 2y$

Solution: *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 2 & -1 \\ 9 & 2 \end{bmatrix}$. *Consider the characteristic equation:*

$$det(A - \lambda I) = det \begin{bmatrix} 2-\lambda & -1 \\ 9 & 2-\lambda \end{bmatrix} = (\lambda - 2)^2 + 9 = 0.$$

Thus $\lambda = 2 \pm 3i$. *Consider* $\lambda = 2 + 3i$, *we seek the e-vector subject to* $(A - (2+3i)I)\vec{u} = 0$. *Solve:*

$$\begin{bmatrix} -3i & -1 \\ 9 & -3i \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad -3iu - v = 0 \quad \Rightarrow \quad v = -3iu, \ u \neq 0 \quad \Rightarrow \quad \vec{u}_1 = \begin{bmatrix} u \\ -3iu \end{bmatrix}$$

We choose $u = 1$ *for convenience and thus find the fundamental complex solution*

$$\vec{x}(t) = e^{(2+3i)t} \begin{bmatrix} 1 \\ -3i \end{bmatrix} = e^{2t}(\cos(3t) + i\sin(3t)) \begin{bmatrix} 1 \\ -3i \end{bmatrix} = e^{2t} \begin{bmatrix} \cos(3t) + i\sin(3t) \\ -3i\cos(3t) + 3\sin(3t) \end{bmatrix}$$

Therefore, using the discussion of the last example, we find fundamental real solutions of the system by selecting real and imaginary parts of the complex solution above:

$$\vec{x}_1(t) = \begin{bmatrix} e^{2t}\cos(3t) \\ 3e^{2t}\sin(3t) \end{bmatrix} \qquad \& \qquad \vec{x}_2(t) = \begin{bmatrix} e^{2t}\sin(3t) \\ -3e^{2t}\cos(3t) \end{bmatrix}.$$

We can assemble the general solution as a linear combination of the fundamental solutions $\vec{x}(t) = c_1 \begin{bmatrix} e^{2t}\cos(3t) \\ 3e^{2t}\sin(3t) \end{bmatrix} + c_2 \begin{bmatrix} e^{2t}\sin(3t) \\ -3e^{2t}\cos(3t) \end{bmatrix}$. Thus the system $x' = 2x - y$ and $y' = 9x + 2y$ has **scalar** solutions $x(t) = c_1 e^{2t}\cos(3t) + c_2 e^{2t}\sin(3t)$ and $y(t) = 3c_1 e^{2t}\sin(3t) - 3c_2 e^{2t}\cos(3t)$. Finally, a fundamental matrix for this problem is given by

$$X = [\vec{x}_1|\vec{x}_2] = \begin{bmatrix} e^{2t}\cos(3t) & e^{2t}\sin(3t) \\ 3e^{2t}\sin(3t) & -3e^{2t}\cos(3t) \end{bmatrix}.$$

Example 4.5.7. Problem: *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & -4 & -1 \\ 0 & 5 & 2 \end{bmatrix}$.

Solution: *Begin by solving the characteristic equation:*

$$\begin{aligned} 0 = det(A - \lambda I) &= det \begin{bmatrix} 2-\lambda & 0 & 0 \\ -1 & -4-\lambda & -1 \\ 0 & 5 & 2-\lambda \end{bmatrix} \\ &= (2-\lambda)\big[(\lambda - 2)(\lambda + 4) + 5\big] \\ &= (2-\lambda)(\lambda - 1)(\lambda + 3). \end{aligned}$$

Thus $\lambda_1 = 1$, $\lambda_2 = 2$ *and* $\lambda_3 = -3$. *We seek* $\vec{u}_1 = [u,v,w]^T$ *such that* $(A - I)\vec{u}_1 = 0$*:*

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & -5 & -1 \\ 0 & 5 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} u = 0 \\ 5v + w = 0 \end{array} \Rightarrow \begin{array}{c} u = 0 \\ w = -5v \end{array} \Rightarrow \vec{u}_1 = v\begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix}.$$

Choose $v = 1$ *for convenience and find* $\vec{u}_1 = [0,1,-5]^T$. *Next, seek* $\vec{u}_2 = [u,v,w]^T$ *such that* $(A - 2I)\vec{u}_2 = 0$*:*

$$\begin{bmatrix} 0 & 0 & 0 \\ -1 & -6 & -1 \\ 0 & 5 & 0 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} -u - 6v - w = 0 \\ v = 0 \end{array} \Rightarrow \begin{array}{c} v = 0 \\ w = -u \end{array} \Rightarrow \vec{u}_2 = u\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Choose $u = 1$ *for convenience and find* $\vec{u}_2 = [1,0,-1]^T$. *Last, seek* $\vec{u}_3 = [u,v,w]^T$ *such that* $(A + 3I)\vec{u}_3 = 0$*:*

$$\begin{bmatrix} 5 & 0 & 0 \\ -1 & -1 & -1 \\ 0 & 5 & 5 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} 5u = 0 \\ 5v + 5w = 0 \end{array} \Rightarrow \begin{array}{c} u = 0 \\ w = -v \end{array} \Rightarrow \vec{u}_3 = v\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Choose $v = 1$ *for convenience and find* $\vec{u}_3 = [0,1,-1]^T$. *The general solution follows:*

$$\boxed{\vec{x}(t) = c_1e^{t}\begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix} + c_2e^{2t}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_3e^{-3t}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.}$$

The fundamental solutions and the fundamental matrix for the system above are given as follows:

$$\vec{x_1}(t) = \begin{bmatrix} 0 \\ e^t \\ -5e^t \end{bmatrix}, \ \vec{x_2}(t) = \begin{bmatrix} e^{2t} \\ 0 \\ -e^{2t} \end{bmatrix}, \ \vec{x_3}(t) = \begin{bmatrix} 0 \\ e^{-3t} \\ -e^{-3t} \end{bmatrix}, \qquad X(t) = \begin{bmatrix} 0 & e^{2t} & 0 \\ e^t & 0 & e^{-3t} \\ -5e^t & -e^{2t} & -e^{-3t} \end{bmatrix}.$$

Example 4.5.8. *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

Solution: *Begin by solving the characteristic equation:*

$$det(A - \lambda I) = det\begin{bmatrix} 2-\lambda & 0 & 0 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 3-\lambda \end{bmatrix} = (2-\lambda)(2-\lambda)(3-\lambda) = 0.$$

Thus $\lambda_1 = 2$, $\lambda_2 = 2$ *and* $\lambda_3 = 3$. *We seek* $\vec{u}_1 = [u,v,w]^T$ *such that* $(A - 2I)\vec{u}_1 = 0$*:*

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} u + w = 0 \\ v \ \ free \end{array} \Rightarrow \begin{array}{c} v \ \ free \\ w = -u \end{array} \Rightarrow \vec{u}_1 = \begin{bmatrix} u \\ v \\ -u \end{bmatrix}.$$

There are two free variables in the solution above and it follows we find two e-vectors. A convenient choice is $u = 1$ *and* $v = 0$ *or* $u = 0$ *and* $v = 1$; $\vec{u}_1 = [1,0,-1]^T$ *and* $\vec{u}_2 = [0,1,0]^T$. *Next, seek* $\vec{u}_3 = [u,v,w]^T$ *such that* $(A - 3I)\vec{u}_3 = 0$

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} u = 0 \\ v = 0 \\ w \ \ free \end{array} \Rightarrow \vec{u}_3 = w\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Choose $w = 1$ for convenience to find $\vec{u}_3 = [0, 0, 1]^T$. The general solution follows:

$$\boxed{\vec{x}(t) = c_1 e^{2t} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 e^{3t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.}$$

The fundamental solutions and the fundamental matrix for the system above are given as follows:

$$\vec{x_1}(t) = \begin{bmatrix} e^{2t} \\ 0 \\ -e^{2t} \end{bmatrix}, \quad \vec{x}_2(t) = \begin{bmatrix} 0 \\ e^{2t} \\ 0 \end{bmatrix}, \quad \vec{x}_3(t) = \begin{bmatrix} 0 \\ 0 \\ e^{3t} \end{bmatrix}, \qquad X(t) = \begin{bmatrix} e^{2t} & 0 & 0 \\ 0 & e^{2t} & 0 \\ -e^{2t} & 0 & e^{3t} \end{bmatrix}.$$

Example 4.5.9. *we seek to solve $\frac{d\vec{x}}{dt} = A\vec{x}$ where $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 3 \end{bmatrix}$.*

Solution: *Begin by solving the characteristic equation:*

$$det \begin{bmatrix} 2-\lambda & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 1 & -1 & 3-\lambda \end{bmatrix} = (2-\lambda)^2(3-\lambda) = 0.$$

Thus $\lambda_1 = 2$, $\lambda_2 = 2$ and $\lambda_3 = 3$. We seek $\vec{u}_1 = [u, v, w]^T$ such that $(A - 2I)\vec{u}_1 = 0$:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{matrix} v = 0 \\ u + w = 0 \end{matrix} \quad \Rightarrow \quad \begin{matrix} v = 0 \\ w = -u \end{matrix} \quad \Rightarrow \quad \vec{u}_1 = \begin{bmatrix} u \\ 0 \\ -u \end{bmatrix}.$$

Choose $u = 1$ to select $\vec{u}_1 = [1, 0, -1]^T$. Next find $\vec{u}_2$ such that $(A - 3I)\vec{u}_2 = 0$

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{matrix} -u + v = 0 \\ -v = 0 \\ w \ \ free \end{matrix} \quad \Rightarrow \quad \begin{matrix} u = 0 \\ v = 0 \\ w \ \ free \end{matrix} \quad \Rightarrow \quad \vec{u}_2 = \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix}.$$

Choose $w = 1$ to find $\vec{u}_2 = [0, 0, 1]^T$. We find two fundamental solutions from the e-vector method:

$$\vec{x}_1(t) = e^{2t} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \qquad \& \qquad \vec{x}_1(t) = e^{3t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We cannot solve the system at this juncture since we are missing the third fundamental solution $\vec{x}_3$. In the next section we will find the missing solution via the generalized e-vector/ matrix exponential method.

Example 4.5.10. *we seek to solve $\frac{d\vec{x}}{dt} = A\vec{x}$ where $A = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$.*

Solution: *Begin by solving the characteristic equation:*

$$det \begin{bmatrix} 7-\lambda & 0 & 0 \\ 0 & 7-\lambda & 0 \\ 0 & 0 & 7-\lambda \end{bmatrix} = (7-\lambda)^3 = 0.$$

Thus $\lambda_1 = 7$, $\lambda_2 = 7$ *and* $\lambda_3 = 7$. *The e-vector equation in this case is easy to solve; since* $A - 7I = 7I - 7I = 0$ *it follows that* $(A - 7I)\vec{u} = 0$ *for all* $\vec{u} \in \mathbb{R}^3$. *Therefore, any nontrivial vector is an eigenvector with e-value 7. A natural choice is* $\vec{u}_1 = [1, 0, 0]^T$, $\vec{u}_2 = [0, 1, 0]^T$ *and* $\vec{u}_3 = [0, 0, 1]^T$. *Thus,*

$$\boxed{\vec{x}(t) = c_1 e^{7t} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 e^{7t} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 e^{7t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = e^{7t} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}.}$$

Example 4.5.11. *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} -2 & 0 & 0 \\ 4 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$.

Solution: *Begin by solving the characteristic equation:*

$$det(A - \lambda I) = det \begin{bmatrix} -2-\lambda & 0 & 0 \\ 4 & -2-\lambda & 0 \\ 1 & 0 & -2-\lambda \end{bmatrix} = -(\lambda + 2)^3 = 0.$$

Thus $\lambda_1 = -2$, $\lambda_2 = -2$ *and* $\lambda_3 = -2$. *We seek* $\vec{u}_1 = [u, v, w]^T$ *such that* $(A + 2I)\vec{u}_1 = 0$*:*

$$\begin{bmatrix} 0 & 0 & 0 \\ 4 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{matrix} u = 0 \\ v \ \ free \\ w \ \ free \end{matrix} \quad \Rightarrow \quad \vec{u}_1 = \begin{bmatrix} 0 \\ v \\ w \end{bmatrix}.$$

Choose $v = 1, w = 0$ *to select* $\vec{u}_1 = [0, 1, 0]^T$ *and* $v = 0, w = 1$ *to select* $\vec{u}_2 = [0, 0, 1]^T$. *Thus we find fundamental solutions:*

$$\vec{x}_1(t) = e^{-2t} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \& \qquad \vec{x}_2(t) = e^{-2t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We cannot solve the system at this juncture since we are missing the third fundamental solution $\vec{x}_3$. *In the next section we will find the missing solution via the generalized e-vector/ matrix exponential method.*

Example 4.5.12. *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 1 \\ 9 & 3 & -4 \end{bmatrix}$.

Solution: *Begin by solving the characteristic equation:*

$$\begin{aligned} det(A - \lambda I) &= det \begin{bmatrix} 2-\lambda & 1 & -1 \\ -3 & -1-\lambda & 1 \\ 9 & 3 & -4-\lambda \end{bmatrix} \\ &= (2-\lambda)[(\lambda+1)(\lambda+4) - 3] - [3(\lambda+4) - 9] - [-9 + 9(\lambda+1)] \\ &= (2-\lambda)[\lambda^2 + 5\lambda + 1] - 3\lambda - 3 - 9\lambda \\ &= -\lambda^3 - 5\lambda^2 - \lambda + 2\lambda^2 + 10\lambda + 2 - 12\lambda - 3 \\ &= -\lambda^3 - 3\lambda^2 - 3\lambda - 1 \\ &= -(\lambda+1)^3 \end{aligned}$$

Thus $\lambda_1 = -1$, $\lambda_2 = -1$ *and* $\lambda_3 = -1$. *We seek* $\vec{u}_1 = [u, v, w]^T$ *such that* $(A + I)\vec{u}_1 = 0$:

$$\begin{bmatrix} 3 & 1 & -1 \\ -3 & 0 & 1 \\ 9 & 3 & -3 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} 3u + v - w = 0 \\ -3u + w = 0 \end{array} \Rightarrow \begin{array}{c} w = 3u \\ v = w - 3u = 0 \end{array} \Rightarrow \vec{u}_1 = \begin{bmatrix} u \\ 0 \\ 3u \end{bmatrix}.$$

Choose $u = 1$ *to select* $\vec{u}_1 = [1, 0, 3]^T$. *We find just one fundamental solution:* $\boxed{\vec{x}_1 = e^{-t}[1, 0, 3]^T}$. *We cannot solve the problem in it's entirety with our current methods. In the section that follows we find the missing pair of solutions.*

Example 4.5.13. *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$.

Solution: *Begin by solving the characteristic equation:*

$$\begin{aligned} det(A - \lambda I) &= det \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & -1 & 1 - \lambda \end{bmatrix} \\ &= -\lambda[\lambda(\lambda - 1) + 1] + 1 \\ &= -\lambda^3 + \lambda^2 - \lambda + 1 \\ &= -\lambda^2(\lambda - 1) - (\lambda - 1) \\ &= (1 - \lambda)(\lambda^2 + 1) \end{aligned}$$

Thus $\lambda_1 = 1$, $\lambda_2 = i$ *and* $\lambda_3 = -i$. *We seek* $\vec{u}_1 = [u, v, w]^T$ *such that* $(A - I)\vec{u}_1 = 0$:

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} -u + v = 0 \\ -v + w = 0 \end{array} \Rightarrow \begin{array}{c} v = u \\ w = v \end{array} \Rightarrow \vec{u}_1 = \begin{bmatrix} u \\ u \\ u \end{bmatrix}.$$

Choose $u = 1$ *thus select* $\vec{u}_1 = [1, 1, 1]^T$. *Now seek* $\vec{u}_2$ *such that* $(A - iI)\vec{u}_2 = 0$

$$\begin{bmatrix} -i & 1 & 0 \\ 0 & -i & 1 \\ 1 & -1 & 1 - i \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{c} v = iu \\ w = iv = i(iu) = -u \\ (i - 1)w = u - v \end{array} \Rightarrow \vec{u}_2 = \begin{bmatrix} u \\ iu \\ -u \end{bmatrix}.$$

Set $u = 1$ *to select the following complex solution:*

$$\vec{x}(t) = e^{it} \begin{bmatrix} 1 \\ i \\ -1 \end{bmatrix} = \begin{bmatrix} e^{it} \\ ie^{it} \\ -e^{it} \end{bmatrix} = \begin{bmatrix} \cos(t) + i\sin(t) \\ i\cos(t) - \sin(t) \\ -\cos(t) - i\sin(t) \end{bmatrix} = \begin{bmatrix} \cos(t) \\ -\sin(t) \\ -\cos(t) \end{bmatrix} + i \begin{bmatrix} \sin(t) \\ \cos(t) \\ -\sin(t) \end{bmatrix}.$$

We select the second and third solutions by taking the real and imaginary parts of the above complex solution; $\vec{x}_2(t) = Re(\vec{x}(t))$ *and* $\vec{x}_3(t) = Im(\vec{x}(t))$. *The general solution follows:*

$$\boxed{\vec{x}(t) = c_1 e^t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} \cos(t) \\ -\sin(t) \\ -\cos(t) \end{bmatrix} + c_3 \begin{bmatrix} \sin(t) \\ \cos(t) \\ -\sin(t) \end{bmatrix}.}$$

The fundamental solution set and fundamental matrix of the example above are simply:

$$\vec{x}_1 = \begin{bmatrix} e^t \\ e^t \\ e^t \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} \cos(t) \\ -\sin(t) \\ -\cos(t) \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} \sin(t) \\ \cos(t) \\ -\sin(t) \end{bmatrix} \qquad \& \qquad X = \begin{bmatrix} e^t & \cos(t) & \sin(t) \\ e^t & -\sin(t) & \cos(t) \\ e^t & -\cos(t) & -\sin(t) \end{bmatrix}$$

4.6 solutions by matrix exponential

Recall the Maclaurin series for the exponential is given by:

$$e^t = \sum_{j=o}^{\infty} \frac{t^j}{j!} = 1 + t + \frac{1}{2}t^2 + \frac{1}{6}t^3 + \cdots$$

This provided the inspiration for the definition given below[9]

Definition 4.6.1. *matrix exponential*

Suppose A is an $n \times n$ matrix then we define the **matrix exponential of** A by:

$$e^A = \sum_{j=0}^{\infty} \frac{A^j}{j!} = I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \cdots$$

Suppose $A = 0$ is the zero matrix. Note that

$$e^0 = I + 0 + \frac{1}{2}0^2 + \cdots = I.$$

Furthermore, as $(-A)^j = (-1)^j A^j$ it follows that $e^{-A} = I - A + \frac{1}{2}A^2 - \frac{1}{6}A^3 + \cdots$. Hence,

$$\begin{aligned} e^A e^{-A} &= \Big(I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \cdots\Big)\Big(I - A + \frac{1}{2}A^2 - \frac{1}{6}A^3 + \cdots\Big) \\ &= I - A + \frac{1}{2}A^2 - \frac{1}{6}A^3 + \cdots + A\Big(I - A + \frac{1}{2}A^2 + \cdots\Big) + \frac{1}{2}A^2\Big(I - A + \cdots\Big) + \frac{1}{6}A^3 I + \cdots \\ &= I + A - A + \frac{1}{2}A^2 - A^2 + \frac{1}{2}A^2 - \frac{1}{6}A^3 + \frac{1}{2}A^3 - \frac{1}{2}A^3 + \frac{1}{6}A^3 + \cdots \\ &= I. \end{aligned}$$

I have only shown the result up to the third-order in A, but you can verify higher orders if you wish. We find an interesting result:

$$\boxed{(e^A)^{-1} = e^{-A} \qquad \Rightarrow \qquad det(e^A) \neq 0 \qquad \Rightarrow \qquad \text{columns of } A \text{ are LI.}}$$

Noncommutativity of matrix multiplication spoils the usual law of exponents. Let's examine how this happens. Suppose A, B are square matrices. Calculate e^{A+B} to second order in A, B:

$$e^{A+B} = I + (A + B) + \frac{1}{2}(A + B)^2 + \cdots = I + A + B + \frac{1}{2}\big(A^2 + AB + BA + B^2\big) + \cdots$$

On the other hand, calculate the product $e^A e^B$ to second order in A, B,

$$e^A e^B = (I + A + \frac{1}{2}A^2 + \cdots)(I + B + \frac{1}{2}B^2 + \cdots) = I + A + B + \frac{1}{2}\big(A^2 + 2AB + B^2\big) + \cdots$$

[9]the concept of an exponential actually extends in much more generality than this, we could derive this from more basic and general principles, but that has little to do with this course so we behave. In addition, the reason the series of matrices below converges is not immediately obvious, see my linear notes for a sketch of the analysis needed here

We find that, to second order, $e^Ae^B - e^{A+B} = \frac{1}{2}(AB - BA)$. Define the **commutator** $[A, B] = AB - BA$ and note (after a short calculation)

$$\boxed{e^Ae^B = e^{A+B+\frac{1}{2}[A,B]+\cdots}}$$

When A, B are **commuting** matrices the commutator $[A, B] = AB - BA = AB - AB = 0$ hence the usual algebra $e^Ae^B = e^{A+B}$ applies. It turns out that the higher-order terms in the boxed formula above can be written as nested-commutators of A and B. This formula is known as the Baker-Campbell-Hausdorff, it is the essential calculation in the theory of matrix Lie groups (which is the math used to formulate important symmetry aspects of modern physics).

Let me pause[10] to give a better proof that $AB = BA$ implies $e^Ae^B = e^{A+B}$. The heart of the argument follows from the fact the binomial theorem holds for $(A + B)^k$ in this context. I seek to prove by mathematical induction on k that $(A + B)^k = \sum_{n=0}^{k} \binom{k}{n} A^{k-n}B^n$. Note $k = 1$ is clearly true as $\binom{1}{0} = \binom{1}{1} = 1$ and $(A + B)^1 = A + B$. Assume inductively the binomial theorem holds for k and seek to prove $k + 1$ true:

$$\begin{aligned}
(A+B)^{k+1} &= (A+B)^k(A+B) \\
&= \left(\sum_{n=0}^{k}\binom{k}{n}A^{k-n}B^n\right)(A+B) \qquad : \text{by induction hypothesis} \\
&= \sum_{n=0}^{k}\binom{k}{n}A^{k-n}B^nA + \sum_{n=0}^{k}\binom{k}{n}A^{k-n}B^nB \\
&= \sum_{n=0}^{k}\binom{k}{n}A^{k-n}AB^n + \sum_{n=0}^{k}\binom{k}{n}A^{k-n}B^{n+1} \qquad : AB = BA \text{ implies } B^nA = AB^n \\
&= \sum_{n=0}^{k}\binom{k}{n}A^{k+1-n}B^n + \sum_{n=0}^{k}\binom{k}{n}A^{k-n}B^{n+1}
\end{aligned}$$

Continuing,

$$\begin{aligned}
(A+B)^{k+1} &= A^{k+1} + \sum_{n=1}^{k}\binom{k}{n}A^{k+1-n}B^n + \sum_{n=0}^{k-1}\binom{k}{n}A^{k-n}B^{n+1} + B^{k+1} \\
&= A^{k+1} + \sum_{n=1}^{k}\binom{k}{n}A^{k+1-n}B^n + \sum_{n=1}^{k}\binom{k}{n-1}A^{k+1-n}B^n + B^{k+1} \\
&= A^{k+1} + \sum_{n=1}^{k}\left[\binom{k}{n} + \binom{k}{n-1}\right]A^{k+1-n}B^n + B^{k+1} \\
&= A^{k+1} + \sum_{n=1}^{k}\binom{k+1}{n}A^{k+1-n}B^n + B^{k+1} \qquad : \text{by Pascal's Triangle} \\
&= \sum_{n=0}^{k+1}\binom{k+1}{n}A^{k+1-n}B^n
\end{aligned}$$

[10] you may skip ahead if you are not interested in how to make arguments precise, in fact, even this argument has gaps, but I include it to give the reader some idea about what is missing when we resort to $+\cdots$-style induction

Which completes the induction step and we find by mathematical induction the binomial theorem for commuting matrices holds for all $k \in \mathbb{N}$. Consider the matrix expontial formula in light of the binomial theorem, also recall $\binom{k+1}{n} = \frac{k!}{n!(k-n)!}$,

$$\begin{aligned} e^{A+B} &= \sum_{k=0}^{\infty} \frac{1}{k!}(A+B)^k \\ &= \sum_{k=0}^{\infty}\sum_{n=0}^{k} \frac{1}{k!}\frac{k!}{n!(k-n)!}A^{k-n}B^n \\ &= \sum_{k=0}^{\infty}\sum_{n=0}^{k} \frac{1}{n!}\frac{1}{(k-n)!}A^{k-n}B^n \end{aligned} \tag{4.2}$$

On the other hand, if we compute the product of e^A with e^B we find:

$$e^A e^B = \sum_{j=0}^{\infty} \frac{1}{j!}A^j \sum_{n=0}^{\infty} \frac{1}{n!}B^n = \sum_{j=0}^{\infty}\sum_{n=0}^{\infty} \frac{1}{n!}\frac{1}{j!}A^j B^n$$

It follows[11] that $e^A e^B = e^{A+B}$. We use this result implicitly in much of what follows in this section.

Suppose A is a constant $n \times n$ matrix. Calculate[12]

$$\begin{aligned} \tfrac{d}{dt}\big[exp(tA)\big] &= \tfrac{d}{dt}\Big[\sum_{k=0}^{\infty}\tfrac{1}{k!}t^kA^k\Big] && \text{defn. of matrix exponential} \\ &= \sum_{k=0}^{\infty}\tfrac{d}{dt}\big[\tfrac{1}{k!}t^kA^k\big] && \text{since matrix exp. uniformly conv.} \\ &= \sum_{k=0}^{\infty}\tfrac{k}{k!}t^{k-1}A^k && A^k \text{ constant and } \tfrac{d}{dt}(t^k) = kt^{k-1} \\ &= A\sum_{k=1}^{\infty}\tfrac{1}{(k-1)!}t^{k-1}A^{k-1} && \text{since } k! = k(k-1)! \text{ and } A^k = AA^{k-1}. \\ &= Aexp(tA) && \text{defn. of matrix exponential.} \end{aligned}$$

I suspect the following argument is easier to follow:

$$\begin{aligned} \tfrac{d}{dt}(exp(tA)) &= \tfrac{d}{dt}(I + tA + \tfrac{1}{2}t^2A^2 + \tfrac{1}{3!}t^3A^3 + \cdots) \\ &= \tfrac{d}{dt}(I) + \tfrac{d}{dt}(tA) + \tfrac{1}{2}\tfrac{d}{dt}(t^2A^2) + \tfrac{1}{3\cdot 2}\tfrac{d}{dt}(t^3A^3) + \cdots \\ &= A + tA^2 + \tfrac{1}{2}t^2A^3 + \cdots \\ &= A(I + tA + \tfrac{1}{2}t^2A^2 + \cdots) \\ &= Aexp(tA). \end{aligned}$$

□

Whichever notation you prefer, the calculation above completes the proof of the following central theorem for this section:

[11] after some analytical arguments beyond this course; what is missing is an explicit examination of the infinite limits at play here, the doubly infinite limits seem to reach the same terms but the structure of the sums differs

[12] the term-by-term differentiation theorem for power series extends to a matrix power series, the proof of this involves real analysis

Theorem 4.6.2.

Suppose $A \in \mathbb{R}^{n\times n}$. The matrix exponential e^{tA} gives a fundamental matrix for $\frac{d\vec{x}}{dt} = A\vec{x}$.

Proof: we have already shown that (1.) e^{tA} is a solution matrix ($\frac{d}{dt}[e^{tA}] = Ae^{tA}$) and (2.) $(e^{tA})^{-1} = e^{-tA}$ thus the columns of e^{tA} are LI. $\square$

It follows that the general solution of $\frac{d\vec{x}}{dt} = A\vec{x}$ is simply $\vec{x}(t) = e^{tA}\vec{c}$ where $\vec{c} = [c_1, c_2, \dots, c_n]^T$ determines the initial conditions of the solution. In theory this is a great formula, we've solved most of the problems we set-out to solve. However, more careful examination reveals this result is much like the result from calculus; any continuous function is integrable. Ok, so f continuous on an interval I implies F exists on I and $F' = f$, but... how do you actually calculate the antiderivative F? It's possible in principle, but in practice the computation may fall outside the computation scope of the techniques covered in calculus[13].

Example 4.6.3. *Suppose $x' = x, y' = 2y, z' = 3z$ then in matrix form we have:*

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The coefficient matrix is diagonal which makes the k-th power particularly easy to calculate,

$$A^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 3^k \end{bmatrix}$$

$$\Rightarrow \quad exp(tA) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 3^k \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{\infty} \frac{t^k}{k!}1^k & 0 & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{t^k}{k!}2^k & 0 \\ 0 & 0 & \sum_{k=0}^{\infty} \frac{t^k}{k!}3^k \end{bmatrix}$$

$$\Rightarrow \quad exp(tA) = \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{3t} \end{bmatrix}$$

Thus we find three solutions to $x' = Ax$,

$$x_1(t) = \begin{bmatrix} e^t \\ 0 \\ 0 \end{bmatrix} \qquad x_2(t) = \begin{bmatrix} 0 \\ e^{2t} \\ 0 \end{bmatrix} \qquad x_3(t) = \begin{bmatrix} 0 \\ 0 \\ e^{3t} \end{bmatrix}$$

In turn these vector solutions amount to the solutions $x = e^t, y = 0, z = 0$ or $x = 0, y = e^{2t}, z = 0$ or $x = 0, y = 0, z = e^{3t}$. It is easy to check these solutions.

Of course the example above is very special. In order to unravel the mystery of just how to calculate the matrix exponential for less trivial matrices we return to the construction of the previous section.

[13] for example, $\int \frac{\sin(x)dx}{x}$ or $\int e^{-x^2}dx$ are known to by incalculable in terms of elementary functions

Let's see what happens when we calculate $e^{tA}\vec{u}$ for $\vec{u}$ and e-vector with e-value λ.

$$\begin{aligned}
e^{tA}\vec{u} &= e^{t(A-\lambda I+\lambda I)}\vec{u} \qquad : \text{ added zero anticipating use of } (A-\lambda I)\vec{u}=0, \\
&= e^{t\lambda I + t(A-\lambda I)}\vec{u} \\
&= e^{t\lambda I}e^{t(A-\lambda I)}\vec{u} \qquad : \text{ noted that } t\lambda I \text{ commutes with } t(A-\lambda I), \\
&= e^{t\lambda}Ie^{t(A-\lambda I)}\vec{u} \qquad : \text{ a short exercise shows } e^{t\lambda I} = e^{t\lambda}I. \\
&= e^{t\lambda}\left(I + t(A-\lambda I) + \frac{t^2}{2}(A-\lambda I)^2 + \cdots\right)\vec{u} \\
&= e^{t\lambda}\big(I\vec{u} + t(A-\lambda I)\vec{u} + \frac{t^2}{2}(A-\lambda I)^2\vec{u} + \cdots\big) \\
&= e^{t\lambda}\vec{u} \qquad : \text{ as it was given } (A-\lambda I)\vec{u}=0 \text{ hence all but the first term vanishes.}
\end{aligned}$$

The fact that this is a a solution of $\vec{x}\,' = A\vec{x}$ was already known to us, however, it is nice to see it arise from the matrix exponential. Moreover the calculation above reveals the central formula that guides the technique of this section. The **magic formula**. For any square matrix and possibly constant λ we find:

$$\boxed{e^{tA} = e^{t\lambda}\left(I + t(A-\lambda I) + \frac{t^2}{2}(A-\lambda I)^2 + \cdots\right) = e^{t\lambda}\sum_{k=0}^{\infty}\frac{t^k}{k!}(A-\lambda I)^k.}$$

When we choose λ as an e-value and multiply this formula by the corresponding e-vector then this infinite series truncates nicely to reveal $e^{\lambda t}\vec{u}$. It follows that we should define vectors which truncate the series at higher order, this is the natural next step:

Definition 4.6.4. *generalized eigenvectors and chains of generalized e-vectors*

Given an eigenvalue λ a nonzero vector $\vec{u}$ such that $(A-\lambda I)^p\vec{u}=0$ is called an **generalized eigenvector of order** p with eigenvalue λ. If $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p$ are nonzero vectors such that $(A-\lambda I)\vec{u}_j = \vec{u}_{j-1}$ for $j = 2, 3, \dots, p$ and $\vec{u}_1$ is an e-vector with e-value λ then we say $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ forms a **chain of generalized e-vectors of length** p.

In the notation of the definition above, it is true that $\vec{u}_k$ is a generalized e-vector of order k with e-value λ. Let's examine $k=2$,

$$(A-\lambda I)\vec{u}_2 = \vec{u}_1 \qquad \Rightarrow \qquad (A-\lambda I)^2\vec{u}_2 = (A-\lambda I)\vec{u}_1 = 0.$$

Then suppose inductively the claim is true for k which means $(A-\lambda I)^k\vec{u}_k = 0$, consider $k+1$

$$(A-\lambda I)\vec{u}_{k+1} = \vec{u}_k \qquad \Rightarrow \qquad (A-\lambda I)^{k+1}\vec{u}_{k+1} = (A-\lambda I)^k\vec{u}_k = 0.$$

Hence, in terms of the notation in the definition above, we have shown by mathematical induction that $\vec{u}_k$ is a generalized e-vector of order k with e-value λ.

I do not mean to claim this is true for all $k \in \mathbb{N}$. In practice for an $n \times n$ matrix we cannot find a chain longer than length n. However, up to that bound such chains are possible for an arbitrary matrix.

Example 4.6.5. *The matrices below are in* **Jordan form** *which means the vectors* $e_1 = [1,0,0,0,0]^T$ *etc...* $e_5 = [0,0,0,0,1]^T$ *are (generalized)-e-vectors:*

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \qquad \& \qquad B = \begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{bmatrix}$$

You can easily calculate $(A-2I)e_1 = 0$, $(A-2I)e_2 = e_1$, $(A-2I)e_3 = e_2$ *or* $(A-3I)e_4 = 0$, $(A-3I)e_5 = e_4$. *On the other hand,* $(B-4I)e_1 = 0$, $(B-4I)e_2 = e_1$ *and* $(B-5I)e_3 = 0$ *and* $(B-6I)e_4 = 0$, $(B-6I)e_5 = 0$. *The matrix* B *needs only one generalized e-vector whereas the matrix* A *has* 3 *generalized e-vectors.*

Let's examine why chains are nice for the magic formula:

Example 4.6.6. Problem: *Suppose* A *is a* 3×3 *matrix with a chain of generalized e-vector* $\vec{u}_1, \vec{u}_2, \vec{u}_3$ *with respect to e-value* $\lambda = 2$. *Solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *in view of these facts.*

Solution: *we are given* $(A-2I)\vec{u}_1 = 0$ *and* $(A-2I)\vec{u}_2 = \vec{u}_1$ *and* $(A-2I)\vec{u}_3 = \vec{u}_2$. *It is easily shown that* $(A-2I)^2\vec{u}_2 = 0$ *and* $(A-2I)^3\vec{u}_3 = 0$. *It is also possible to prove* $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ *is a LI set. Apply the magic formula with* $\lambda = 2$ *to derive the following results:*

1. $\vec{x}_1(t) = e^{tA}\vec{u}_1 = e^{2t}\vec{u}_1$ *(we've already shown this in general earlier in this section)*
2. $\vec{x}_2(t) = e^{tA}\vec{u}_2 = e^{2t}\big(I\vec{u}_2 + t(A-2I)\vec{u}_2 + \frac{t^2}{2}(A-2I)^2\vec{u}_2 + \cdots\big) = e^{2t}(\vec{u}_2 + t\vec{u}_1)$.
3. *note that* $(A-2I)^2\vec{u}_3 = (A-2I)\vec{u}_2 = \vec{u}_1$ *hence:*

$$\vec{x}_3(t) = e^{tA}\vec{u}_3 = e^{2t}\big(I\vec{u}_3 + t(A-2I)\vec{u}_3 + \frac{t^2}{2}(A-2I)^2\vec{u}_3 + \cdots\big) = e^{2t}(\vec{u}_3 + t\vec{u}_2 + \frac{t^2}{2}\vec{u}_1).$$

Therefore, $\boxed{\vec{x}(t) = c_1e^{2t}\vec{u}_1 + c_2e^{2t}(\vec{u}_2 + t\vec{u}_1) + c_3e^{2t}(\vec{u}_3 + t\vec{u}_2 + \frac{t^2}{2}\vec{u}_1)}$ *is the general solution.*

Perhaps it is interesting to calculate $e^{tA}[\vec{u}_1|\vec{u}_2|\vec{u}_3]$ in view of the calculations in the example above. Observe:

$$e^{tA}[\vec{u}_1|\vec{u}_2|\vec{u}_3] = [e^{tA}\vec{u}_1|e^{tA}\vec{u}_2|e^{tA}\vec{u}_3] = e^{2t}\left[\vec{u}_1 \middle| \vec{u}_2 + t\vec{u}_1 \middle| \vec{u}_3 + t\vec{u}_2 + \frac{t^2}{2}\vec{u}_1\right].$$

I suppose we could say more about this formula, but let's get back on task: we seek to complete the solution of the unsolved problems of the previous section. It is our hope that we can find generalized e-vector solutions to complete the fundamental solution sets in Examples 4.5.3, 4.5.9,4.5.11 and 4.5.12.

Example 4.6.7. Problem: (returning to Example 4.5.3) *solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 3 & 1 \\ -4 & -1 \end{bmatrix}$

Solution: *we found* $\lambda_1 = 1$ *and* $\lambda_2 = 1$ *and a single e-vector* $\vec{u}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. *Now seek a generalized e-vector* $\vec{u}_2 = [u,v]^T$ *such that* $(A-I)\vec{u}_2 = \vec{u}_1$,

$$\begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \Rightarrow \quad 2u + v = 1 \quad \Rightarrow \quad v = 1 - 2u, \;\; u \neq 0 \quad \Rightarrow \quad \vec{u}_1 = \begin{bmatrix} u \\ 1-2u \end{bmatrix}$$

We choose $u = 0$ for convenience and thus find $\vec{u}_2 = [0,1]^T$ hence the fundamental solution

$$\vec{x}_2(t) = e^{tA}\vec{u}_2 = e^t(I + t(A-I) + \cdots)\vec{u}_2 = e^t(\vec{u}_2 + t\vec{u}_1) = e^t \begin{bmatrix} t \\ 1-2t \end{bmatrix}.$$

Therefore, we find $\boxed{\vec{x}(t) = c_1 e^t \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 e^t \begin{bmatrix} t \\ 1-2t \end{bmatrix}}.$

Example 4.6.8. Problem: (returning to Example 4.5.9) *solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 3 \end{bmatrix}$.

Solution: *we found $\lambda_1 = 2$, $\lambda_2 = 2$ and $\lambda_3 = 3$ and we also found e-vector $\vec{u}_1 = [1,0,-1]^T$ with e-value 2 and e-vector $\vec{u}_2 = [0,0,1]^T$. Seek $\vec{u}_3$ such that $(A-2I)\vec{u}_3 = \vec{u}_1$ since we are missing a solution paired with $\lambda_2 = 2$.*

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \Rightarrow \begin{matrix} v = 1 \\ u - 1 + w = -1 \end{matrix} \Rightarrow \begin{matrix} v = 1 \\ w = -u \end{matrix} \Rightarrow \vec{u}_3 = \begin{bmatrix} u \\ 1 \\ -u \end{bmatrix}.$$

Choose $u = 0$ to select $\vec{u}_3 = [0,1,0]^T$. It follows from the magic formula that $\vec{x}_3(t) = e^{tA}\vec{u}_3 = e^{2t}(\vec{u}_3 + t\vec{u}_1)$. Hence, the general solution is

$$\boxed{\vec{x}(t) = c_1 e^{2t} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_2 e^{3t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + c_3 e^{2t} \begin{bmatrix} t \\ 1 \\ -t \end{bmatrix}}.$$

Once more we found a generalized e-vector of order two to complete the solution set and find $\vec{x}_3$ in the example above. You might notice that had we replaced the choice $u = 0$ in both of the last examples with some nonzero u then we would have added a copy of $\vec{x}_1$ to the generalized e-vector solution. This is permissable since the sum of solutions to the system $\vec{x}\,' = A\vec{x}$ is once more a solution. This freedom works hand-in-hand with the ambiguity of the generalized e-vector problem.

Example 4.6.9. Problem: (returning to Example 4.5.11) *we seek to solve* $\frac{d\vec{x}}{dt} = A\vec{x}$ *where* $A = \begin{bmatrix} -2 & 0 & 0 \\ 4 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$.

Solution: *We already found $\lambda_1 = -2$, $\lambda_2 = -2$ and $\lambda_3 = -2$ and a pair of e-vectors $\vec{u}_1 = [0,1,0]^T$ and $v = 0, w = 1$ to select $\vec{u}_2 = [0,0,1]^T$. We face a dilemma, should we look for a chain that ends with $\vec{u}_1 = [0,1,0]^T$ or $\vec{u}_2 = [0,0,1]^T$? Generally it may not be possible to do either. Thus, we set aside the chain condition and instead look for directly for solutions of $(A+2I)^2\vec{u}_3 = 0$.*

$$(A+2I)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 4 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 4 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Since we seek $\vec{u}_3$ which forms a LI set with $\vec{u}_1, \vec{u}_2$ it is natural to choose $\vec{u}_3 = [1,0,0]^T$. Calculate,

$$\begin{aligned} \vec{x}_3(t) = e^{tA}\vec{u}_3 &= e^{-2t}(I\vec{u}_3 + t(A+2I)\vec{u}_3 + \frac{t^2}{2}(A+2I)^2\vec{u}_3 + \cdots) \qquad (4.3) \\ &= e^{-2t}\left[\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 & 0 & 0 \\ 4 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right] \\ &= e^{-2t}\begin{bmatrix} 1 \\ 4t \\ t \end{bmatrix} \end{aligned}$$

Thus we find the general solution:

$$\boxed{\vec{x}(t) = c_1e^{-2t}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_2e^{-2t}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + c_3e^{-2t}\begin{bmatrix} 1 \\ 4t \\ t \end{bmatrix}}.$$

I leave the complete discussion of the chains in the subtle case above for the second course on linear algebra. See Insel Spence and Friedberg's *Linear Algebra* text for an accessible treatment aimed at advanced undergraduates.

Example 4.6.10. Problem: (returning to Example 4.5.12) *we seek to solve $\frac{d\vec{x}}{dt} = A\vec{x}$ where* $A = \begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 1 \\ 9 & 3 & -4 \end{bmatrix}$.

Solution: *we found $\lambda_1 = -1$, $\lambda_2 = -1$ and $\lambda_3 = -1$ and a single e-vector $\vec{u}_1 = [1,0,3]^T$. Seek $\vec{u}_2$ such that $(A+I)\vec{u}_2 = \vec{u}_1$,*

$$\begin{bmatrix} 3 & 1 & -1 \\ -3 & 0 & 1 \\ 9 & 3 & -3 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \Rightarrow \begin{matrix} 3u+v-w=1 \\ -3u+w=0 \end{matrix} \Rightarrow \begin{matrix} w=3u \\ v=w-3u+1 \end{matrix} \Rightarrow \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

where we set $u = 0$ for convenience. Continuing, we seek $\vec{u}_3$ where $(A+I)\vec{u}_3 = \vec{u}_2$,

$$\begin{bmatrix} 3 & 1 & -1 \\ -3 & 0 & 1 \\ 9 & 3 & -3 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} 3u+v-w=0 \\ -3u+w=1 \end{matrix} \Rightarrow \begin{matrix} w=1+3u \\ v=w-3u \end{matrix} \Rightarrow \begin{matrix} w=1+3u \\ v=1 \end{matrix}$$

Choose $u = 0$ to select $\vec{u}_3 = [0,1,1]^T$. Given the algebra we've completed we know that

$$(A+I)\vec{u}_1 = (A+I)^2\vec{u}_2 = (A+I)^3\vec{u}_3 = 0, \quad (A+I)\vec{u}_2 = \vec{u}_1, \quad (A+I)\vec{u}_3 = \vec{u}_2, \quad (A+I)^2\vec{u}_3 = \vec{u}_1$$

These identities paired with the magic formula with $\lambda = -1$ yield:

$$e^{tA}\vec{u}_1 = e^{-t}\vec{u}_1 \qquad \& \qquad e^{tA}\vec{u}_2 = e^{-t}(\vec{u}_2 + t\vec{u}_1) \qquad \& \qquad e^{tA}\vec{u}_3 = e^{-t}(\vec{u}_3 + t\vec{u}_2 + \frac{t^2}{2}\vec{u}_1)$$

Therefore, we find general solution:

$$\boxed{\vec{x}(t) = c_1e^{-t}\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + c_2e^{-t}\begin{bmatrix} t \\ 1 \\ 3t \end{bmatrix} + c_3e^{-t}\begin{bmatrix} \frac{t^2}{2} \\ 1+t \\ 1+\frac{3t^2}{2} \end{bmatrix}}.$$

The method we've illustrated extends naturally to the case of repeated complex e-values where there are insufficient e-vectors to form the general solution.

Example 4.6.11. Problem: *Suppose A is a 6×6 matrix with a chain of generalized e-vector $\vec{u}_1, \vec{u}_2, \vec{u}_3$ with respect to e-value $\lambda = 2+i$. Solve $\frac{d\vec{x}}{dt} = A\vec{x}$ in view of these facts.*

Solution: *we are given $(A-(2+i)I)\vec{u}_1 = 0$ and $(A-(2+i)I)\vec{u}_2 = \vec{u}_1$ and $(A-(2+i)I)\vec{u}_3 = \vec{u}_2$. It is easily shown that $(A-(2+i)I)^2\vec{u}_2 = 0$ and $(A-(2+i)I)^3\vec{u}_3 = 0$. It is also possible to prove $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is a LI set. Apply the magic formula with $\lambda = (2+i)$ to derive the following results:*

1. $\vec{x}_1(t) = e^{tA}\vec{u}_1 = e^{(2+i)t}\vec{u}_1$ *(we've already shown this in general earlier in this section)*
2. $\vec{x}_2(t) = e^{tA}\vec{u}_2 = e^{(2+i)t}\big(I\vec{u}_2 + t(A-(2+i)I)\vec{u}_2 + \frac{t^2}{2}(A-(2+i)I)^2\vec{u}_2 + \cdots\big) = e^{(2+i)t}(\vec{u}_2 + t\vec{u}_1)$.
3. *note that* $(A-(2+i)I)^2\vec{u}_3 = (A-(2+i)I)\vec{u}_2 = \vec{u}_1$ *hence:*

$$\vec{x}_3(t) = e^{tA}\vec{u}_3 = e^{(2+i)t}\big(I\vec{u}_3 + t(A-(2+i)I)\vec{u}_3 + \frac{t^2}{2}(A-(2+i)I)^2\vec{u}_3 + \cdots\big) = e^{(2+i)t}(\vec{u}_3 + t\vec{u}_2 + \frac{t^2}{2}\vec{u}_1).$$

The solutions $\vec{x}_1(t)$, $\vec{x}_2(t)$ and $\vec{x}_3(t)$ are complex-valued solutions. To find the real solutions we select the real and imaginary parts to form the fundamental solution set

$$\{Re(\vec{x}_1), Im(\vec{x}_1), Re(\vec{x}_2), Im(\vec{x}_2), Re(\vec{x}_3), Im(\vec{x}_3)\}$$

I leave the explicit formulas to the reader, it is very similar to the case we treated in the last section for the complex e-vector problem.

Suppose A is idempotent or order k then $A^{k-1} \neq I$ and $A^k = I$. In this case the matrix exponential simplifies:

$$e^{tA} = I + tA + \frac{t^2}{2}A^2 + \cdots + \frac{t^{k-1}}{(k-1)!}A^{k-1} + \left(\frac{t^k}{k!} + \frac{t^{k+1}}{(k+1)!} + \cdots\right)I$$

However, $\frac{t^k}{k!} + \frac{t^{k+1}}{(k+1)!} + \cdots = e^t - 1 - t - \frac{t^2}{2} - \cdots - \frac{t^{k-1}}{(k-1)!}$ hence we can calculate e^{tA} nicely in such a case. On the other hand, if the matrix A is nilpotent of order k then $A^{k-1} \neq 0$ and $A^k = 0$. Once again, the matrix exponential simplifies:

$$e^{tA} = I + tA + \frac{t^2}{2}A^2 + \cdots + \frac{t^{k-1}}{(k-1)!}A^{k-1}$$

Therefore, if A is nilpotent then we can calculate the matrix exponential directly without too much trouble... of course this means we can solve $\vec{x}\,' = A\vec{x}$ without use of the generalized e-vector method.

Finally, I conclude this section with a few comments about direct computation via the Cayley Hamilton Theorem (this is proved in an advanced linear algebra course)

Theorem 4.6.12.

If $A \in \mathbb{R}^{n \times n}$ and $p(\lambda) = det(A - \lambda I) = 0$ is the characteristic equation then $p(A) = 0$.

Note that if $p(x) = x^2 + 3$ then $p(A) = A^2 + 3I$.

Example 4.6.13. Problem: solve the system given in Example 4.5.12) *by applying the Cayley Hamilton Theorem to* $A = \begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 1 \\ 9 & 3 & -4 \end{bmatrix}$.

Solution: *we found* $p(\lambda) = -(\lambda - 1)^3 = 0$ *hence* $-(A - I)^3 = 0$. *Consider the magic formula:*

$$e^{tA} = e^t(I + t(A - I) + \frac{t^2}{2}(A - I)^2 + \frac{t^3}{3!}(A - I)^3 + \cdots) = e^t(I + t(A - I) + \frac{t^2}{2}(A - I)^2)$$

Calculate,

$$A - I = \begin{bmatrix} 1 & 1 & -1 \\ -3 & -2 & 1 \\ 9 & 3 & -5 \end{bmatrix} \qquad \& \qquad (A - I)^2 = \begin{bmatrix} -11 & -4 & 5 \\ 12 & 4 & -4 \\ -45 & -12 & 19 \end{bmatrix}$$

Therefore,

$$e^{tA} = e^t \begin{bmatrix} 1 + t - \frac{11t^2}{2} & t - 2t^2 & -t + \frac{5t^2}{2} \\ -3t + 6t^2 & 1 - 2t + 2t^2 & t - 2t^2 \\ 9t - \frac{45t^2}{2} & 3t - 6t^2 & 1 - 5t + \frac{19t^2}{2} \end{bmatrix}$$

The general solution is given by $\vec{x}(t) = e^{tA}\vec{c}$.

There are certainly additional short-cuts and deeper understanding that stem from a working knowledge of full-fledged linear algebra, but, I hope I have shown you more than enough in these notes to solve any constant-coefficient system $\vec{x}\,' = A\vec{x}$. It turns out there are always enough generalized e-vectors to complete the solution. The existence of the basis made of generalized e-vectors (called a **Jordan basis**) is a deep theorem of linear algebra. It is often, sadly, omitted from undergraduate linear algebra texts. The pair of examples below illustrate some of the geometry behind the calculations of this section.

Example 4.6.14. *Consider for example, the system*

$$x' = x + y, \qquad y' = 3x - y$$

We can write this as the matrix problem

$$\underbrace{\begin{bmatrix} x' \\ y' \end{bmatrix}}_{d\vec{x}/dt} = \underbrace{\begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\vec{x}}$$

It is easily calculated that A *has eigenvalue* $\lambda_1 = -2$ *with e-vector* $\vec{u}_1 = (-1, 3)$ *and* $\lambda_2 = 2$ *with e-vectors* $\vec{u}_2 = (1, 1)$. *The general solution of* $d\vec{x}/dt = A\vec{x}$ *is thus*

$$\vec{x}(t) = c_1 e^{-2t} \begin{bmatrix} -1 \\ 3 \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -c_1 e^{-2t} + c_2 e^{2t} \\ 3c_1 e^{-2t} + c_2 e^{2t} \end{bmatrix}$$

So, the **scalar solutions** *are simply* $\boxed{x(t) = -c_1 e^{-2t} + c_2 e^{2t}}$ *and* $\boxed{y(t) = 3c_1 e^{-2t} + c_2 e^{2t}}$.

Thus far I have simply told you how to solve the system $d\vec{x}/dt = A\vec{x}$ with e-vectors, it is interesting to see what this means geometrically. For the sake of simplicity we'll continue to think about the preceding example. In it's given form the DEqn is **coupled** which means the equations for the derivatives of the dependent variables x, y cannot be solved one at a time. We have to solve both at once. In the next example I solve the same problem we just solved but this time using a change of variables approach.

Example 4.6.15. *Suppose we change variables using the diagonalization idea: introduce new variables* $\bar{x}, \bar{y}$ *by* $P(\bar{x}, \bar{y}) = (x, y)$ *where* $P = [\vec{u}_1|\vec{u}_2]$*. Note* $(\bar{x}, \bar{y}) = P^{-1}(x, y)$*. We can diagonalize* A *by the similarity transformation by* P*;* $D = P^{-1}AP$ *where* $Diag(D) = (-2, 2)$*. Note that* $A = PDP^{-1}$ *hence* $d\vec{x}/dt = A\vec{x} = PDP^{-1}\vec{x}$*. Multiply both sides by* P^{-1}*:*

$$P^{-1}\frac{d\vec{x}}{dt} = P^{-1}PDP^{-1}\vec{x} \quad \Rightarrow \quad \frac{d(P^{-1}\vec{x})}{dt} = D(P^{-1}\vec{x}).$$

You might not recognize it but the equation above is decoupled. In particular, using the notation $(\bar{x}, \bar{y}) = P^{-1}(x, y)$ *we read from the matrix equation above that*

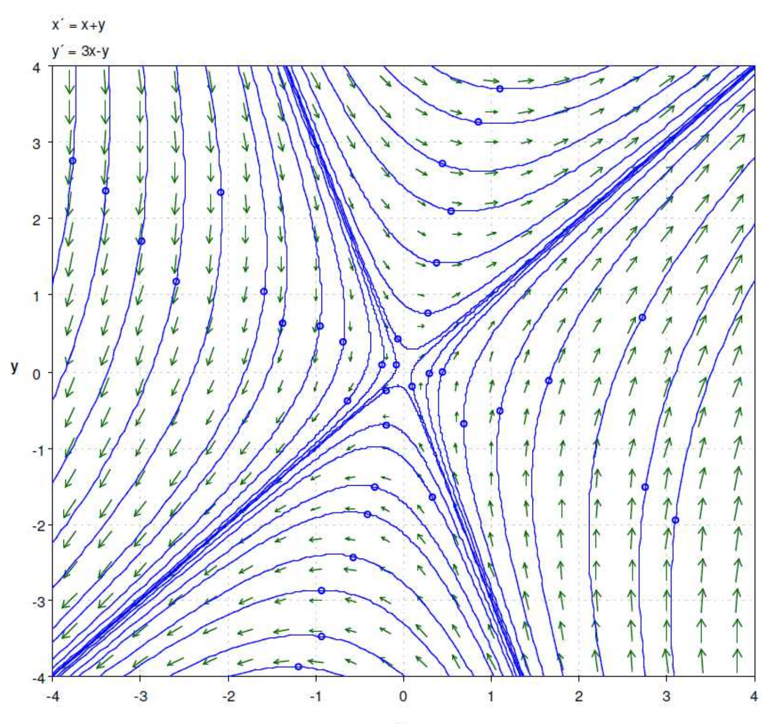
$$\frac{d\bar{x}}{dt} = -2\bar{x}, \qquad \frac{d\bar{y}}{dt} = 2\bar{y}.$$

Separation of variables and a little algebra yields that $\bar{x}(t) = c_1e^{-2t}$ *and* $\bar{y}(t) = c_2e^{2t}$*. Finally, to find the solution back in the original coordinate system we multiply* $P^{-1}\vec{x} = (c_1e^{-2t}, c_2e^{2t})$ *by* P *to isolate* $\vec{x}$*,*

$$\vec{x}(t) = \begin{bmatrix} -1 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} c_1e^{-2t} \\ c_2e^{2t} \end{bmatrix} = \begin{bmatrix} -c_1e^{-2t} + c_2e^{2t} \\ 3c_1e^{-2t} + c_2e^{2t} \end{bmatrix}.$$

This is the same solution we found in the last example. Usually linear algebra texts present this solution because it shows more interesting linear algebra, however, from a pragmatic viewpoint the first method is clearly faster.

Finally, we can better appreciate the solutions we found if we plot the direction field $(x', y') = (x+y, 3x-y)$ via the "pplane" tool in Matlab. I have clicked on the plot to show a few representative trajectories (solutions):



4.7 nonhomogeneous problem

Theorem 4.7.1.

The nonhomogeneous case $\vec{x}\,' = A\vec{x} + \vec{f}$ the general solution is $\vec{x}(t) = X(t)\vec{c} + \vec{x}_p(t)$ where $\vec{c}$ is a vector of constants, X is a fundamental matrix for the corresponding homogeneous system and $\vec{x}_p$ is a particular solution to the nonhomogeneous system. We can calculate $\vec{x}_p(t) = X(t)\int X^{-1}\vec{f}dt$.

Proof: suppose that $\vec{x}_p = X\vec{v}$ for X a fundamental matrix of $\vec{x}\,' = A\vec{x}$ and some vector of unknown functions $\vec{v}$. We seek conditions on $\vec{v}$ which make $\vec{x}_p$ satisfy $\vec{x}_p\,' = A\vec{x}_p + \vec{f}$. Consider,

$$(\vec{x}_p)' = (X\vec{v})' = X'\vec{v} + X\vec{v}' = AX\vec{v} + X\vec{v}'$$

But, $\vec{x}_p\,' = A\vec{X}_p + \vec{f} = AX\vec{v} + \vec{f}$ hence

$$X\tfrac{d\vec{v}}{dt} = \vec{f} \quad \Rightarrow \quad \tfrac{d\vec{v}}{dt} = X^{-1}\vec{f}$$

Integrate to find $\vec{v} = \int X^{-1}\vec{f}dt$ therefore $x_p(t) = X(t)\int X^{-1}\vec{f}dt$. $\square$

If you ever work through variation of parameters for higher order ODEqns then you should appreciate the calculation above. In fact, we can derive n-th order variation of parameters from converting the n-th order ODE by reduction of order to a system of n first order linear ODEs. You can show that the so-called Wronskian of the fundamental solution set is precisely the determinant of the fundamental matrix for the system $\vec{x}\,' = A\vec{x}$ where A is the companion matrix.

Example 4.7.2. Problem: *Suppose that* $A = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}$ *and* $\vec{f} = \begin{bmatrix} e^t \\ e^{-t} \end{bmatrix}$, *find the general solution of the nonhomogenous DEqn* $\vec{x}\,' = A\vec{x} + \vec{f}$.

Solution: *you can easily show* $\vec{x}\,' = A\vec{x}$ *has fundamental matrix* $X = \begin{bmatrix} 1 & e^{4t} \\ -3 & e^{4t} \end{bmatrix}$. *Use variation of parameters for systems of ODEs to constuct* $\vec{x}_p$. *First calculate the inverse of the fundamental matrix, for a* 2×2 *we know a formula:*

$$X^{-1}(t) = \tfrac{1}{e^{4t}-(-3)e^{4t}}\begin{bmatrix} e^{4t} & -e^{4t} \\ 3 & 1 \end{bmatrix} = \tfrac{1}{4}\begin{bmatrix} 1 & -1 \\ 3e^{-4t} & e^{-4t} \end{bmatrix}$$

Thus,

$$\begin{aligned} x_p(t) = X(t)\int \tfrac{1}{4}\begin{bmatrix} 1 & -1 \\ 3e^{-4t} & e^{-4t} \end{bmatrix}\begin{bmatrix} e^t \\ e^{-t} \end{bmatrix}dt &= \tfrac{1}{4}X(t)\int \begin{bmatrix} e^t - e^{-t} \\ 3e^{-3t} + e^{-5t} \end{bmatrix}dt \\ &= \tfrac{1}{4}\begin{bmatrix} 1 & e^{4t} \\ -3 & e^{4t} \end{bmatrix}\begin{bmatrix} e^t + e^{-t} \\ -e^{-3t} - \frac{1}{5}e^{-5t} \end{bmatrix} \\ &= \tfrac{1}{4}\begin{bmatrix} 1(e^t + e^{-t}) + e^{4t}(-e^{-3t} - \frac{1}{5}e^{-5t}) \\ -3(e^t + e^{-t}) + e^{4t}(-e^{-3t} - \frac{1}{5}e^{-5t}) \end{bmatrix} \\ &= \tfrac{1}{4}\begin{bmatrix} e^t + e^{-t} - e^t - \frac{1}{5}e^{-t} \\ -3e^t - 3e^{-t} - e^t - \frac{1}{5}e^{-t} \end{bmatrix} \\ &= \tfrac{1}{4}\begin{bmatrix} \frac{4}{5}e^{-t} \\ -4e^t - \frac{16}{5}e^{-t} \end{bmatrix} \end{aligned}$$

Therefore, the general solution is

$$\vec{x}(t) = c_1 \begin{bmatrix} 1 \\ -3 \end{bmatrix} + c_2 e^{4t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \tfrac{1}{5} \begin{bmatrix} e^{-t} \\ -5e^{t} - 4e^{-t} \end{bmatrix}.$$

The general scalar solutions implicit within the general vector solution $\vec{x}(t) = [x(t), y(t)]^T$ *are*

$$x(t) = c_1 + c_2 e^{4t} + \tfrac{1}{5}e^{-t} \qquad y(t) = -3c_1 + c_2 e^{4t} - e^{t} - \tfrac{4}{5}e^{-t}.$$

I might ask you to solve a 3×3 system in the homework. The calculation is nearly the same as the preceding example with the small inconvenience that finding the inverse of a 3×3 requires some calculation.

Remark 4.7.3.

You might wonder how would you solve a system of ODEs $x' = Ax$ such that the coefficients A_{ij} are not constant. The theory we've discussed holds true with appropriate modification of the interval of applicability. In the constant coefficient case $I = \mathbb{R}$ so we have had no need to discuss it. In order to solve non-constant coefficient problems we will need to find a method to solve the homogeneous problem to locate the fundamental matrix. Once that task is accomplished the technique of this section applies to solve any associated nonhomogeneous problem.

4.8 practice problems

PP 236 Solve $LI_1'' + R_1 I_1' + \frac{1}{C}(I_1' - I_2') = 0$ and $R_2 I_2' + \frac{1}{C}(I_2 - I_1) = \mathcal{E}'(t)$ given that $\mathcal{E}(t) = 10\cos(2t)$, $L = 1$ and $C = 1$ and $R_1 = 2$ and $R_2 = 3$ (in volts, seconds, Henries and Farads). These differential equations stem from the circuit pictured below:

PP 237 Suppose $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$. Show that $e^D = \begin{bmatrix} e^{d_1} & 0 \\ 0 & e^{d_2} \end{bmatrix}$.

PP 238 Suppose $\frac{dx}{dt} = x + 4y$ and $\frac{dy}{dt} = x + y$. Find the general real solution via the e-vector method.

PP 239 Suppose $\frac{dx}{dt} = 2x + y$ and $\frac{dy}{dt} = 2y$. Find the general real solution via the generalized e-vector method.

PP 240 Suppose $\frac{dx}{dt} = 4x - 3y$ and $\frac{dy}{dt} = 3x + 4y$. Find the general real solution via the e-vector method.

PP 241 Suppose $\frac{dx}{dt} = x + 4y + e^{6t}$ and $\frac{dy}{dt} = x + y + 3$. Find the solution with $x(0) = 0$ and $y(0) = 0$. Please use matrix arguments (do not solve by the operator method, instead, use variation of parameters for systems)

PP 242 Suppose X is a fundamental matrix for $\frac{d\vec{x}}{dt} = A\vec{x}$. Suppose B is a square matrix with $det(B) \neq 0$. Show that XB is a fundamental matrix for $\frac{d\vec{x}}{dt} = A\vec{x}$.

PP 243 Calculate e^{tA} for $A = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}$. *(Problem 238 should help)*

PP 244 Use the Cayley Hamilton Theorem to calculate e^{tA} for $A = \begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 1 \\ 9 & 3 & -4 \end{bmatrix}$. The Cayley Hamilton Theorem simply states that a matrix solves it's own characteristic equation; that is, if $p(\lambda) = 0$ is the characteristic equation then $p(A) = 0$. For example, if $p(\lambda) = (\lambda + 2)^3 = 0$ then $(A + 2I)^3 = 0$. The proof of this theorem is easy in the diagonalizable case, however the general proof requires ideas about invariant subspaces often not covered in the undergraduate course on linear algebra.

you may use technology to aid with the matrix calculations in the next three problems. That said, you don't really need it for these in my view

PP 245 Suppose $\frac{dx}{dt} = 5x - 6y - 6z$, $\frac{dy}{dt} = -x + 4y + 2z$ and $\frac{dz}{dt} = 3x - 6y - 4z$. Find the general real solution via the e-vector method.

PP 246 Suppose $\frac{dx}{dt} = 5x - 5y - 5z$, $\frac{dy}{dt} = -x + 4y + 2z$ and $\frac{dz}{dt} = 3x - 5y - 3z$. Find the general real solution via the e-vector method.

PP 247 Suppose $\frac{dx}{dt} = 3x + y$, $\frac{dy}{dt} = 3y + z$ and $\frac{dz}{dt} = 3z$. Find the general real solution via the generalized e-vector method.

PP 248 To solve $\frac{d\vec{x}}{dt} = A\vec{x}$ in the case $A = \begin{bmatrix} -3 & 0 & -3 \\ 1 & -2 & 3 \\ 1 & 0 & 1 \end{bmatrix}$ by the following calculations:

(a) find the e-values and corresponding e-vectors $\vec{u_1}, \vec{u_2}, \vec{u_3}$. (you may use technology)

(b) construct $P = [\vec{u_1}|\vec{u_2}|\vec{u_3}]$ and calculate $P^{-1}AP$. (you may use technology)

(c) note the solution of $AP\vec{y} = \frac{d}{dt}[P\vec{y}] = P\frac{d\vec{y}}{dt}$ is easily found since multiplying by P^{-1} yields $P^{-1}AP\vec{y} = P^{-1}P\frac{d\vec{y}}{dt} = I\frac{d\vec{y}}{dt} = \frac{d\vec{y}}{dt}$. Solve $P^{-1}AP\vec{y} = \frac{d\vec{y}}{dt}$. (this should be really easy, just solve 3 first order problems, one at a time)

(d) $AP\vec{y} = \frac{d}{dt}[P\vec{y}]$ means $\vec{x} = P\vec{y}$ solves $\frac{d\vec{x}}{dt} = A\vec{x}$. Solve the original system by multiplying the solution from (3.) by P.

The method outlined above is more meaningful in a larger discussion involving coordinate change for linear transformations. The coordinates $\vec{y} = P^{-1}\vec{x}$ are **eigencoordinates**. *A matrix is said to be*

diagonalizable *iff there exists some coordinate change matrix P such that $P^{-1}AP = D$ where D is diagonalizable. Not all matrices are diagonalizable. We've seen this. When there are less than n-LI e-vectors then we cannot build the P-matrix as above and it turns out there is no other way to diagonalize a matrix. On the other hand, the generalized e-vectors always exist and cojugating by P made of generalized e-vectors will place* **any** *matrix in Jordan-form (possibly complex).*

PP 249 Suppose A is an 7×7 matrix with complex e-value $\lambda_1 = 3i$ repeated and a real e-value of $\lambda_2 = 1$ repeated three times. You are given a complex vector $\vec{u}_1 = \vec{a}_1 + i\vec{b}_1$ a second LI complex-vector $\vec{u}_2 = \vec{a}_2 + i\vec{b}_2$ such that

$$(A - 3iI)\vec{u}_1 = 0 \qquad (A - 3iI)\vec{u}_2 = \vec{u}_1.$$

We assume $\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{b}_2$ are all real vectors. Furthermore, you are given $\vec{u}_3, \vec{u}_4, \vec{u}_5$ LI vectors such that

$$(A - I)\vec{u}_3 = 0, \qquad A\vec{u}_4 = \vec{u}_4, \qquad A\vec{u}_5 = \vec{u}_5 + \vec{u}_4$$

Find the general, manifestly real, solution.

PP 250 Suppose a force $F(x) = 3x^4 + 16x^3 + 6x^2 - 72x$ is the net-force on some mass $m = 1$. Newton's Equation is $\ddot{x} = 3x^4 + 16x^3 + 6x^2 - 72x$.

(a) make the substitution $v = \dot{x}$ and write Newton's equation as a system in normal form for x and v.

(b) find all three critical points for the system in (1.). (the potential should factor nicely)

(c) plot the potential plane and phase plane juxtaposed vertically with the potential at the top and the phase plane at the base. Plot several trajectories and include arrows to indicate the direction of physically feasible solutions.

(d) classify each critical point by examining your plot from (3.)

in this context the phase plane is also called the Poincare plane in honor of the mathematician who did much pioneering work in this realm of qualitative analysis. Incidentally, given any autonomous system $\frac{dx}{dt} = g(x, y)$ *and* $\frac{dy}{dt} = f(x, y)$ *we can study the timeless phase plane equation* $\frac{dy}{dx} = \frac{f}{g}$ *to indirectly analyze the solutions to the system. Solutions to the phase plane equation are the Cartesian level curves which are parametrized, with parameter* t*, by the solutions to the system*

PP 251 The Volterra-Lotka equations are a nonlinear system of ODEs which model the population interaction between some prey with population x and predator of population y. For example, $\frac{dx}{dt} = x(3 - y)$ and $\frac{dy}{dt} = y(x - 3)$. This means that when the predator population is over 3 then prey population declines. On the other hand, if the prey population goes beyond 3 then the predator population grows. This competition can lead to a variety of outcomes. Find all the critical points of the system and plot the phase plane via the pplane tool, plot about 20 interesting trajectories. Comment on the stability of the critical points. (you'll need to print this out and attach it to this homework)

PP 252 Show that nontrivial solutions for the cauchy-euler system $t\frac{d\vec{x}}{dt} = A\vec{x}$ of the form $\vec{x}(t) = t^R\vec{u}$ must have R an e-value of A with $\vec{u}$ the corresponding e-vector. Solve $t\frac{d\vec{x}}{dt} = A\vec{x}$ in the case $A = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix}$ for $t > 0$.

PP 253 Difference equations can sometimes be written in the form $\vec{x}_{k+1} = B\vec{x}_k$ where $k = 0, 1, 2, \dots$. It is easy to show that if $\vec{x}_o$ is the given **initial state** of the system then the k-th state is found by $\vec{x}_k = B^k\vec{x}_o$. There is a natural connection with this difference equation and the linear differential equations we have studied. Consider this: for small $\triangle t$,

$$\frac{d\vec{x}}{dt} = A\vec{x} \quad \Rightarrow \quad \frac{\vec{x}(t + \triangle t) - \vec{x}(t)}{\triangle t} \approxeq A\vec{x}(t) \quad \Rightarrow \quad \vec{x}(t + \triangle t) = \vec{x}(t) + \triangle t A\vec{x}(t)$$

Hence, $\vec{x}(t + \triangle t) = (I + \triangle t A)\vec{x}(t)$. Identify that this approximation resembles the difference equation where $\vec{x}(t) = \vec{x}_k$ and $\vec{x}(t + \triangle t) = \vec{x}_{k+1}$ and $B = I + \triangle t A$.

(a) Suppose $\vec{x}_o = [2,0]^T$ is the initial state. Calculate the states up to $k = 10$ for $\vec{x}_{k+1} = B\vec{x}_k$ where $B = \begin{bmatrix} 1.1 & -1 \\ 1 & 1.1 \end{bmatrix}$.

(b) Solve $\frac{d\vec{x}}{dt} = A\vec{x}$ where $A = \begin{bmatrix} 0.1 & -1 \\ 1 & 0.1 \end{bmatrix}$ given the initial condition $\vec{x}(0) = [2,0]^T$.

(c) plot the states from (1.) as dots and the solution from (2.) as a curve on a common xy-plane. Comment on what you see. (what $\triangle t$ did I choose ? How could we make the difference equation more closely replicate the differential equation?)

PP 254 Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Calculate A^2.

PP 255 Let A be as in the previous problem. Suppose $v_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}$.

(a.) calculate Av_1 c

(b.) calculate Av_2

(c.) calculate $A[v_1|v_2]$ (here $[v_1|v_2]$ is the 3×2 matrix made from gluing (aka concatenating) the column vectors v_1 and v_2)

(d.) Does $A[v_1|v_2] = [Av_1|Av_2]$?

PP 256 A square matrix X is invertible iff there exists Y such that $XY = YX = I$ where I is the identity matrix. Moreover, linear algebra reveals that X is invertible iff $det(X) \neq 0$. For a 2×2 matrix $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ we define $det(X) = ad - bc$. Suppose X is invertible and show $X^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. This formula is worth memorizing for future use in two-dimensional problems. Please understand, all I'm asking here is for you to multiply X and my propsed formula for X^{-1} to obtain $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

PP 257 Differentiation of matrices of functions is not hard. Let $X(t) = \begin{bmatrix} e^t & t \\ 1/t & e^{-t} \end{bmatrix}$. Calculate:

(a.) calculate $\frac{dX}{dt}$

(b.) calculate $\frac{dX^{-1}}{dt}$

(c.) simplify $\frac{dX}{dt}X^{-1} + X\frac{dX^{-1}}{dt}$.

(d.) explain the previous part by differentiating $X(t)X^{-1}(t) = I$. Note: the product rule for matrix products is simply $\frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}$.

PP 258 If two masses m_1, m_2 are coupled by a spring and then the whole system is attached to springs between to walls (see figure 1 on page 230 of Ritger & Rose for a related picture) then

$$m_1\ddot{x}_1 = -k_1x_1 + k_2(x_2 - x_1)$$
$$m_2\ddot{x}_2 = -k_2(x_2 - x_1) - k_3x_2.$$

PP 259 Suppose $k_2 = 0$. Find the equations of motion.

PP 260 Suppose $k_1 = k_3 = 0$. Find the equations of motion.

PP 261 Suppose $k_1 = k_3 = 1$ and $k_2 = 2$ with $m_1 = m_2 = 1$. Find the equations of motion.

PP 262 Solve $x' = 7x + 3y$ and $y' = 3x + 7y$ by the eigenvector method.

PP 263 Use the solution of the previous problem to solve $x' = 7x + 3y + 1$ and $y' = 3x + 7y + 2$ subject the initital condition $x(0) = 1$ and $y(0) = 2$.

PP 264 Solve $x' = -3x - 5y$ and $y' = 3x + y$ with $x(0) = 4$ and $y(0) = 0$ by the eigenvector method.

PP 265 An ice tray has tiny holes between each of its three partititions such that the water can flow from one partition to the next. Let x, y, z denote the height of water in the three water troughs. The holes are designed such that the flow rate is proportional to the height of water above the adjacent trough. For example, supposing x and z are the edge troughs whereas y is in the middle we have $\frac{dx}{dt} = k(y - x)$. For simplicity of discussion suppose $k = 1$. Write the corresponding differential equations to find the water-level in the y and z troughs. If initially there is 3.0 cm of water in the x trough and none in the other two troughs then find the height in all three troughs as a function of time t. Discuss the steady state solution, is it reasonable?

PP 266 Let a, b be constants which are some measure of the trust between two nations. Furthermore, let x be the military expenditure of Bobslovakia and let y be the military expenditure of the Leaf Village. Detailed analysis by strategically gifted ninjas reveal that

$$\frac{dx}{dt} = -x + 2y + a$$

$$\frac{dy}{dt} = 4x - 3y + b$$

Analyze possible outcomes for various initial conditions and values of a, b. Consider drawing an ab-plane to explain your solution(s). Is a stable peace without a run-away arms race possible given the analysis thus far?

PP 267 Suppose $(A - \lambda I)\vec{u}_1 = 0$ and $(A - \lambda I)\vec{u}_2 = \vec{u}_1$ where $\lambda = 3 + i\sqrt{2}$ and $\vec{u}_1 = [3 + i, 4 + 2i, 5 + 3i, 6 + 4i]^T$ and $\vec{u_2} = [i, 1, 2, 3 - i]^T$.

(a.) find a pair of complex solutions of $\frac{d\vec{x}}{dt} = A\vec{x}$

(b.) extract four real solutions to write the general real solution (c_1, c_2, c_3, c_4 should be real in this answer)

PP 268 Let $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and let $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Calculate $e^{\theta J}$ where $\theta \in \mathbb{R}$. Express your answer in terms of sine and cosine and relevant matrices.

PP 269 Solve $x' = 2x + y$ and $y' = 2y$ by the method generalized eigenvectors.

PP 270 Introduce variables to reduce

$$y''' + 4y'' + 2y' + 6y = \tan(t)$$

to a system of three first order ODEs in matrix normal form $\frac{d\vec{x}}{dt} = A\vec{x} + \vec{f}$.

PP 271 Introduce variables to reduce

$$y'' + 4ty' + 5y' = 0, \qquad w'' + 9e^{-t}w = 0$$

to a system of four first order ODEs in matrix normal form $\frac{d\vec{x}}{dt} = A\vec{x}$.

PP 272 Linear independence (LI) of vector-valued functions $\{\vec{f_j} : I \subseteq \mathbb{R} \to \mathbb{R}^n \mid j = 1, \dots, k\}$ is defined in the same way as was previously discussed for real-valued functions. In particular, $\{\vec{f_1}, \dots, \vec{f_k}\}$ is LI on $I \subseteq \mathbb{R}$ if $c_1\vec{f_1}(t) + \cdots + c_k\vec{f_k}(t) = 0$ for all $t \in I$ implies $c_1 = 0, \dots, c_k = 0$. We can check LI of n such n-vector-valued functions without any further differentiation; in particular, if $\det[\vec{f_1}(t)|\cdots|\vec{f_n}(t)] \neq 0$ for all $t \in I \subseteq \mathbb{R}$ then $\{\vec{f_1}(t), \dots, \vec{f_n}(t)\}$ is LI on I. Show the following sets of vector-valued functions are LI on $\mathbb{R}$. (notice, my notation is that $(a, b) = [a, b]^T$, in other words, each of the expressions below has lists of column vectors.

(a.) $\{\,(e^t, e^t),\ (e^t, -e^t)\,\}$

(b.) $\{\,(\cos(t), -\sin(t)),\ (\sin(t), \cos(t))\,\}$,

(c.) $\{\,e^t\vec{u}_1,\ ,\ e^t(\vec{u}_2 + t\vec{u}_1)\ , e^t(\vec{u}_3 + t\vec{u}_2 + \frac{t^2}{2}\vec{u}_3)\,\}$ given $\vec{u}_1 = (1, 0, 0), \vec{u}_2 = (0, 1, 1), \vec{u}_3 = (1, 1, 1)$.

PP 273 (Cook 5.1)(problem 13 of section 4.9 in Zill) Solve:

$$2\dot{x} - 5x + \dot{y} = e^t$$

$$2\dot{x} - x + \dot{y} = 5e^t$$

PP 274 (Cook 5.1)(problem 7 of section 7.6 in Zill) Solve:

$$\ddot{x} + x - y = 0,$$

$$\ddot{y} + y - x = 0,$$

subject the initial conditions $x(0) = 0, \dot{x}(0) = -2$ and $y(0) = 0, \dot{y}(0) = 1$. (you could use the technique of section 4.9 or that of 7.6, either method should be a profitable exercise)

PP 275 (matrix multiplication) work problem 6 of Appendix II in Zill (page APP-18)

PP 276 Solve, via the eigenvector technique,

$$\frac{dx}{dt} = 5x - y$$
$$\frac{dy}{dt} = -x + 5y.$$

PP 277 Plot the direction field of the system given in previous Problem using pplane. Plot a few solutions. Can you see the e-vectors' geometric significance? Include a print-out of your investigation.

PP 278 Solve, via the complex eigenvector technique,

$$\frac{dx}{dt} = 4x + 2y$$
$$\frac{dy}{dt} = -x + 2y.$$

PP 279 Plot the direction field of the system given in the previous problem. Plot a few solutions. Can you see the e-vectors' geometric significance? Include a print-out of your investigation.

PP 280 Solve $x' = 7x + 3y + 4z,\ y' = 6x + 2y,\ z' = 5z$ by the eigenvector method.

PP 281 Use technology to find e-values and e-vectors for each of the matrices below. If possible, use the solutions of $\frac{d\vec{x}}{dt} = A\vec{x}$ derived from e-vectors to write the general solution of $\frac{d\vec{x}}{dt} = A\vec{x}$. If not possible, explain why.

(a.) $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}.$

(b.) $A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$

(c.) $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{bmatrix}$

(d.) $A = \begin{bmatrix} -1 & -3 & -9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{bmatrix}.$

(e.) $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$

PP 282 Find fundamental matrices for each of the systems given in the previous half dozen problems where reasonable.

PP 283 Suppose $\vec{v}$ is an eigevector with eigenvalue λ for the real matrix A. Show A^2 also has e-vector $\vec{v}$. What is the e-value for $\vec{v}$ with respect to A^2.

PP 284 Write down the magic formula for the matrix exponential.

PP 285 Suppose A is a 3×3 matrix with nonzero vectors $\vec{u}, \vec{v}, \vec{w}$ such that

$$A\vec{u} = 3\vec{u}, \qquad (A - 3I)\vec{v} = \vec{u}, \qquad A\vec{w} = 0.$$

Write the general solution of $\frac{d\vec{x}}{dt} = A\vec{x}$ in terms of the given vectors.

PP 286 Suppose $(A - \lambda I)\vec{u}_1 = 0$ and $(A - \lambda I)\vec{u}_2 = \vec{u}_1$ where $\lambda = 3 + i\sqrt{2}$ and $\vec{u}_1 = [3 + i, 4 + 2i, 5 + 3i, 6 + 4i]^T$ and $\vec{u_2} = [i, 1, 2, 3 - i]^T$.

(a.) find a pair of complex solutions of $\frac{d\vec{x}}{dt} = A\vec{x}$

(b.) extract four real solutions to write the general real solution (c_1, c_2, c_3, c_4 should be real in this answer)

PP 287 Let $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and let $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Calculate $e^{\theta J}$ where $\theta \in \mathbb{R}$. Express your answer in terms of sine and cosine and relevant matrices.

PP 288 Solve $x' = 2x + y$ and $y' = 2y$ by the method generalized eigenvectors.

PP 289 Show why $\frac{d}{dt}e^{tA} = Ae^{tA}$. Is this enough to show e^{tA} is a fundamental solution matrix? If not, say what else we need to know about the matrix exponential.

PP 290 Show $\vec{x}(t) = e^{tA}\vec{x}_o$ is a solution to $\frac{d\vec{x}}{dt} = A\vec{x}$ with $\vec{x}(0) = \vec{x}_o$. In this sense, the matrix exponential generates the solution of the system of ODEs with coefficient matrix A.

PP 291 (matrix inverse of 2 × 2) Suppose $X(t) = \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}$. Find $X^{-1}(t)$. (use the nice formula in Example 5.2.7 of Cook)

PP 292 work out problem 15 of section 8.3.2 in Zill. That is, solve $\frac{d\vec{x}}{dt} = A\vec{x} + \vec{f}$ where $A = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}$ and $\vec{f}(t) = \begin{bmatrix} e^t \\ -e^t \end{bmatrix}$

PP 293 work out problem 21 of section 8.3.2 in Zill. That is, solve $\frac{d\vec{x}}{dt} = A\vec{x} + \vec{f}$ where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $\vec{f}(t) = \begin{bmatrix} \sec t \\ 0 \end{bmatrix}$

PP 294 Consider the differential equation $y'' - 2y' + y = 0$. I think we can all solve this one. Let $x_1 = y, x_2 = y'$. Let A be the companion matrix which stems from the reduction of order just listed. Solve $\frac{d\vec{x}}{dt} = A\vec{x}$ by translating the fundamental solution set $\{y_1, y_2\} = \{e^t, te^t\}$ into the corresponding fundamental solution set $\{\vec{x}_1, \vec{x}_2\}$. Let $\vec{u}_1 = e^{-t}\vec{x}_1$ and $\vec{u}_2 = e^{-t}\vec{x}_2$. Solve the following equations:

$$(A - I)\vec{u}_1 = \vec{a} \qquad (A - I)\vec{u}_2 = \vec{b}.$$

In other words, find $\vec{a}, \vec{b}$ explicitly. Comment on which of the fundamental solutions to $\{\vec{x}_1, \vec{x}_2\}$ was an eigensolution.

PP 295 Suppose A has n-LI e-vectors and hence we can write the general solution for $\frac{d\vec{x}}{dt} = A\vec{x}$ as a linear combination

$$\vec{x} = c_1 e^{\lambda_1 t} + \cdots + c_n e^{\lambda_n t}$$

Solve $\frac{d\vec{x}}{dt} = A^k\vec{x}$ where $k \in \mathbb{N}$.

PP 296 If $A^T = A$ then we say A is a symmetric matrix. A rather deep theorem of linear algebra states that a symmetric matrix has real eigenvalues and it is possible to select n-LI eigenvectors $\{\vec{u}_1, \dots, \vec{u}_n\}$ for which $A\vec{u}_j = \lambda_j\vec{u}_j$ and $\vec{u}_i \cdot \vec{u}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ for all $i, j \in \mathbb{N}_n$. It follows that $P = [\vec{u}_1|\cdots|\vec{u}_n]$ has $P^TP = I$ which means $P^{-1} = P^T$. This means, if we're studying a system of differential equations $\frac{d\vec{x}}{dt} = A\vec{x}$ with $A^T = A$ we can change coordinates to $\vec{y} = P^T\vec{x}$ and in that new $\vec{y}$-coordinate system the differential equation is simply:

$$\frac{dy_1}{dt} = \lambda_1 y_1, \ \dots, \ \frac{dy_n}{dt} = \lambda_n y_n. \qquad \star.$$

This system is said to be **uncoupled** and it's really the most trivial sort of system you can come across; we can solve each equation in the uncoupled system without knowledge of the remaining variables. Consider $\vec{x} = \langle x, y, z\rangle$ and the differential equation $\frac{d\vec{x}}{dt} = A\vec{x}$ where

$A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$. Find an orthonormal eigenbasis for A and use it to change coordinates on the given system. Verify the claim $\star$ in the context of A. Use the notation $\vec{y} = \langle \bar{x}, \bar{y}, \bar{z} \rangle$, so $y_1 = \bar{x}$ etc..

PP 297 Consider the solution-set of $4xy + 4xz + 4yz = 1$. Change to the barred-coordinates $\bar{x}, \bar{y}, \bar{z}$ you discovered in the previous problem. Which Quadric surface is this?

PP 298 The Cayley Hamilton Theorem states that a matrix will solve its own characteristic equation. For example, if $P(x) = x^3 + I$ then $P(A) = A^3 + I = 0$. For this A, calculate e^{tA} in terms of A. Recall, as you should know, $e^{tA} = I + tA + \cdots = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n$.

PP 299 Solve $x' = -x - 4y$ and $y' = 8x + 11y$ using matrix methods.

PP 300 Solve $x' = -7x - 6y$ and $y' = 15x + 11y$ using matrix methods.

PP 301 Suppose $A = \begin{bmatrix} 2 & 5 \\ 0 & 2 \end{bmatrix}$. Calculate e^{tA}.
Also, solve $\frac{d\vec{r}}{dt} = A\vec{r}$ given that $\vec{r}(0) = (1, 2)$.

PP 302 Consider A is a 3×3 matrix for which there exist nonzero vectors v_1, v_2, v_3 such that:

$$Av_1 = 10v_1, \quad Av_2 = 10v_2, \quad Av_3 = 10v_3 + v_1$$

derive the general solution for $\frac{d\vec{r}}{dt} = A\vec{r}$ with appropriate arguments based on the matrix exponential.

PP 303 Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ let $B = I + A$ where $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Also set $M = \begin{bmatrix} 8 & 5 & 9 \\ 6 & 3 & 0 \\ 7 & 0 & 0 \end{bmatrix}$

(a.) Calculate AB and calculate BA. (this doesn't usually happen)

(b.) We say $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ in $\mathbb{R}^3$. Calculate Me_1 and Me_2 then check that $M \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = [Me_1 | Me_2]$. *This ought to illustrate the column-by-column multiplication rule in the sense that* $M[e_1|e_2] = [Me_1|Me_2]$. *Recall this was important for us as we analyzed how the solution matrix gives us a matrix where each column is itself a solution*

PP 304 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 3 & 0 \end{bmatrix}$. Calculate the following items for A,

(a.) show the eigenvalues of A are $\lambda_1 = -3$, $\lambda_2 = -1$ and ,$\lambda_3 = 6$

(b.) find eigenvectors $\vec{u}_1, \vec{u}_2, \vec{u}_3$ with eigenvalues $\lambda_1 = -3$, $\lambda_2 = -1$ and ,$\lambda_3 = 6$ respective. Normalize the eigenvectors so that each has length one.

(c.) show $\vec{u}_i \bullet \vec{u}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$.

(d.) Let $P = [\vec{u}_1|\vec{u}_2|\vec{u}_3]$ and show that $P^T P = I$ (this shows that $P^{-1} = P^T$ and that P is what is known as an **orthgonal matrix**)

(e.) Calculate $P^T AP$. You should get something really pretty.

Remark: the problem above illustrates the real spectral theorem which implies that a symmetric matrix has an orthonormal eigenbasis and eigenvalues which are all real

PP 305 Find the general solution of $\frac{d\vec{r}}{dt} = A\vec{r}$ where A is was given in the previous problem.

PP 306 Let $x' = 2x - 3y$ and $y' = 3x + 2y$. Find the general real solution via the techinque of eigenvectors and/or generalized eigenvectors. In addition, set-up the solution to $x' = 2x - 3y + f_1$ and $y' = 3x + 2y + f_2$ via the method of variation of parameters for systems. *hint: I think this one requires calculation of a complex eigenvector*

PP 307 Let $x' = 3x - 18y$ and $y' = 2x - 9y$. Find the general real solution via the techinque of eigenvectors and/or generalized eigenvectors. Then solve the initial value problem for the given system of DEqns with initial data $x(0) = 1$ and $y(0) = 0$. *hint: I believe this problem will require you to find one eigenvector and one generalized eigenvector, both with the same eigevalue*

PP 308 Let I be the 2×2 identity matrix and let

$$K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Prove that $e^{tK} = \cosh(t)I + \sinh(t)K$. Is $\frac{d\vec{r}}{dt} = K\vec{r}$ a system of differential equations obtained by reduction of order ? If so, do the solutions you found in e^{tK} coincide logically with those you find by directly solving the corresponding 2-nd order problem ?

PP 309 Suppose A is a 4×4 matrix with nonzero real vectors $\vec{u}_1, \vec{u}_2, \vec{u}_3$ and $\vec{u}_4$ for which:

$$A\vec{u_1} = 3\vec{u}_1, \quad (A - 3I)\vec{u_2} = \vec{u_3}, \quad (A - 3I)\vec{u}_3 = \vec{u_1}, \qquad A\vec{u}_4 = 0$$

Find the general solution to $\frac{d\vec{r}}{dt} = A\vec{r}$. Do not assume it fits a pattern. You need to THINK.

Chapter 5

energy analysis and the phase plane approach

This chapter collects our thoughts on how to use energy to study problems in Newtonian mechanics. In particular we explain how to plot possible motions in the Poincare plane ($x, \dot{x}$ plane, or the one-dimensional tangent bundle if you're interested). A simple method allows us to create plots in the Poincare plane from corresponding data for the plot of the potential energy function. Nonconservative examples can be included as modifications of corresponding conservative systems.

All of this said, there are mathematical techniques which extend past physical examples. We begin by discussing such generic features. In particular, the nature of critical points for autonomous linear ODEs have structure which is revealed from the spectrum (list of eigenvalues from smallest to largest) of the coefficient matrix. In fact, such observations are easily made for n-dimensional problems. Of course our graphical methods are mainly of use for two-dimensional problems. We discuss almost linear systems and some of the deeper results due to Poincare for breadth. We omit discussion of Liapunov exponents, however the interested reader would be well-advised to study that topic along side what is discussed here (chapter 10 of Ritger & Rose has much to add to these notes).

Time-permitting we may exhibit the linearization of a non-linear system of ODEs and study how successful our approximation of the system is relative to the numerical data exhibited via the pplane tool. We also may find time to study the method of characteristics as presented in Zachmanoglou and Thoe and some of the deeper symmetry methods which are describe in Peter Hydon's text or Brian Cantwell's text on symmetries in differerential equations.

5.1 phase plane and stability

This section concerns largely qualitative analysis for systems of ODEs. We know from the existence theorems the solutions to a system of ODEs can be unique and will exist given continuity of the coefficient matrix which defines the system. However, certain points where the derivatives are all zero are places where interesting things tend to happen to the solution set. Often many solutions merge at such a **critical point**.

Definition 5.1.1. *critical point for a system of ODEs in normal form*

If the system of ODEs $\frac{d\vec{x}}{dt} = F(\vec{x}, t)$ has a solution $\vec{x}$ for which t_o has $\frac{d\vec{x}}{dt}(t_o) = 0$ then $\vec{x}(t_o)$ is called a **critical point** of the system.

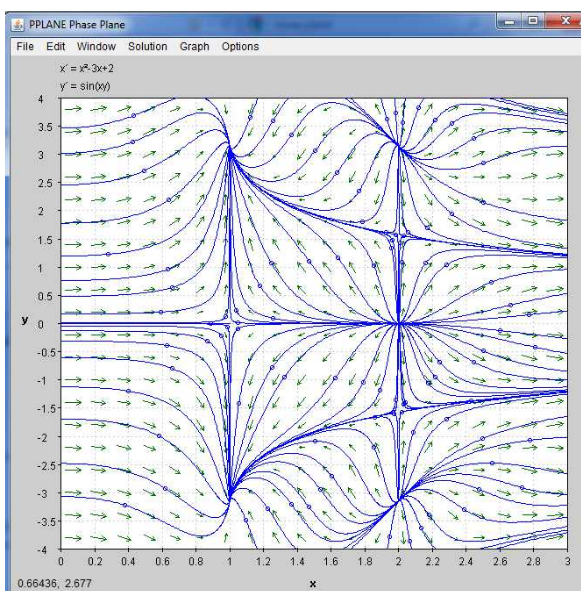
There are two major questions that concern us: (1.) where are the critical point(s) for a given system of ODEs ? (2.) do solutions near a given critical point tend to stay near the point or flow far away ? Let us begin by studying a system of two **autonomous ODEs**

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

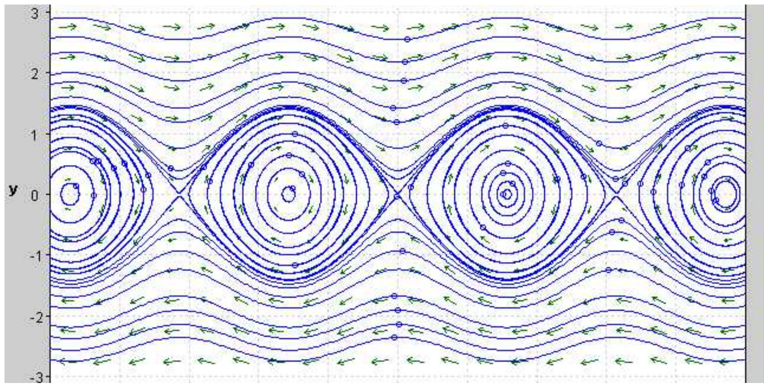
The location of critical points becomes an algebra problem: the system above has a critical point wherever both $f(x_o, y_o) = 0$ and $g(x_o, y_o) = 0$.

Example 5.1.2. Problem: *find critical points of the system* $\frac{dx}{dt} = x^2 - 3x + 2$, $\frac{dy}{dt} = \sin(xy)$.

Solution: *a critical point must simultaneously solve* $x^2 - 3x + 2 = 0$ *and* $\sin(xy) = 0$. *The polynomial equation factors to yield* $(x-1)(x-2) = 0$ *hence we require the point to have either* $x = 1$ *or* $x = 2$. *If* $x = 1$ *then* $\sin(y) = 0$ *hence* $y = n\pi$ *for* $n \in \mathbb{Z}$. *It follows that* $(1, n\pi)$ *is a critical point for each* $n \in \mathbb{Z}$. *Likewise, if* $x = 2$ *then* $\sin(2y) = 0$ *hence* $2y = k\pi$ *for* $k \in \mathbb{Z}$ *hence* $y = k\pi/2$. *It follows that* $(2, k\pi/2)$ *is a critical point for each* $k \in \mathbb{Z}$.



The plot above was prepared with the *pplane* tool which you can find online. You can study the plot and you'll spot the critical points with ease. If you look more closely then you'll see that some of the critical points have solutions which flow into the point whereas others have solutions which flow out of the point. If all the solutions flow into the point then we say the point is **stable** or **asymptotically stable**. Otherwise, if some solutions flow away from the point without bound then the point is said to be **unstable**. I will not attempt to give careful descriptions of these terms here. There is another type of stable point. Let me illustrate it by example. The plot below shows sample solutions for the system $\frac{dx}{dt} = y$ and $\frac{dy}{dt} = \sin(2x)$. The points where $y = 0$ and $x = n\pi/2$ for some $n \in \mathbb{Z}$ are critical points.



These critical points are **stable centers**. Obviously I used pplane to create the plot above, but another method is known and has deep physical significance for problems such as the one illustrated above. The method I discuss next is known as the **energy method**, I focus on a class of problems which directly stem from a well-known physical problem.

Consider a mass m under the influence of a conservative force $F = -dU/dx$. Note:

$$ma = F \quad \Rightarrow \quad m\frac{d^2x}{dt^2} = -\frac{dU}{dx} \quad \Rightarrow \quad m\frac{dx}{dt}\frac{d^2x}{dt^2} = -\frac{dx}{dt}\frac{dU}{dx} \quad \Rightarrow \quad m\frac{dx}{dt}\frac{dv}{dt} = -\frac{dx}{dt}\frac{dU}{dx}$$

However, $v\frac{dv}{dt} = \frac{d}{dt}[\frac{1}{2}v^2]$ and $\frac{dx}{dt}\frac{dU}{dx} = \frac{dU}{dt}$ hence,

$$m\frac{d}{dt}[\frac{1}{2}v^2] = -\frac{dU}{dt} \quad \Rightarrow \quad \frac{d}{dt}\left[\frac{1}{2}mv^2 + U\right] = 0$$

In particular, we find that if x, v are solutions of $ma = F$ then the associated **energy function:**

$$\boxed{E(x,v) = \frac{1}{2}mv^2 + U(x)}$$

is constant along solutions of Newton's Second Law. Furthermore, consider $m\frac{d^2x}{dt^2} - F(x) = 0$ as a second order ODE. We can reduce it two a system of two ODEs in normal form by the standard substitution: $v = \frac{dx}{dt}$. Using velocity as an independent coordinate gives:

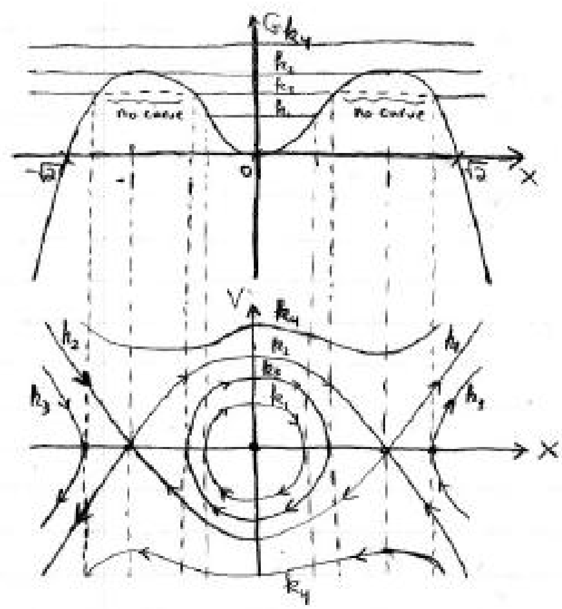
$$\boxed{\frac{dx}{dt} = v \qquad \& \qquad \frac{dv}{dt} = \frac{F}{m}}$$

Critical points of this system occur wherever both $v = 0$ and $F = 0$ since $m > 0$ by physical assumptions. Given our calculations concerning energy the solutions to this system must somehow parametrize the energy level curves as they appear in the xv-plane. This xv-plane is called the

phase plane or the **Poincare plane** in honor of the mathematician who pioneered these concepts in the early 20-th century. Read Chapter 5 of Nagel Saff and Snider for a brief introduction to the concept of chaos and how the Poincare plane gave examples which inspired many mathematicians to work on the problem over the century that followed (chaos is still an active math research area).

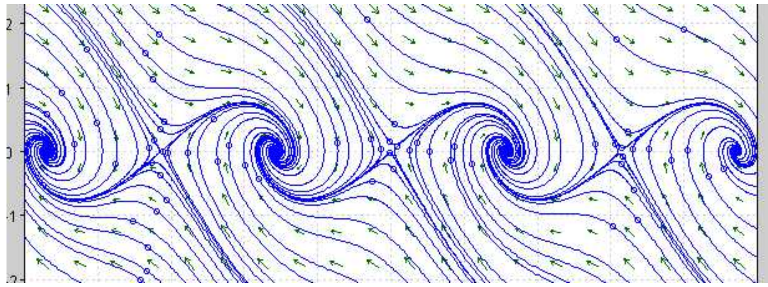
Think further about the critical points of $\frac{dx}{dt} = v$ & $\frac{dv}{dt} = \frac{F}{m}$. Recall we assumed F was conservative hence there exists a **potential energy** function U such that $F = -\frac{dU}{dx}$. This means the condition $F = 0$ gives $\frac{dU}{dx} = 0$. **Ah HA !** this means that the critical points of the phase plane solutions must be on the x-axis (where $v = 0$) at points where the potential energy U has critical points in the xU-plane. Here I am contrasting the concept of critical point of a system with critical point ala calculus I. The xU-plane is called the **potential plane**.

The analysis of the last paragraph means that we can use the potential energy diagram to create the phase plane trajectories. This is closely tied to the specific mathematics of the energy function. Let us observe for a particular energy $E_o = \frac{1}{2}mv^2 + U(x)$ we cannot have motions where $U(x) < E_o$ since the kinetic energy $\frac{1}{2}mv^2 \geq 0$. Moreover, points where $E = U$ are points where $v = 0$ and these correspond to points where the motion either turns around or is resting.



In the plot above the top-graph is the **potential plane plot** whereas the lower plot is the corresponding **phase plane plot**. The point $(0, 0)$ is a stable center in the phase plane whereas $(\pm 1, 0)$ are unstable critical points. The trajectories in the phase plane are constructed such that the critical points match-up and the direction of all trajectories with $v > 0$ flow right whereas those with $v < 0$ flow left since $\frac{dx}{dt} = v$. Also, if $E > U$ at a critical point of U then the corresponding trajectory will have a horizontal tangent since $\frac{dv}{dt} = 0$ at such points. These rules force you to draw essentially the same pattern plotted above.

All of the discussion above concerns the conservative case. In retrospect you should see the example $\frac{dx}{dt} = y$ and $\frac{dy}{dt} = \sin(2x)$ is the phase plane DEqn corresponding to $m = 1$ with $F = \sin(2x)$. If we add a friction force $F_f = -v$ then $\frac{dv}{dt} = \sin(2x) - v$ is Newton's equation and we would study the system $\frac{dx}{dt} = y$ and $\frac{dy}{dt} = \sin(2x) - y$. The energy $E = \frac{1}{2}v^2 - \frac{1}{2}\cos(2x)$ is not conserved in this case. I will not work out the explicit details of such analysis here, but perhaps you will find the contrast of the pplane plot below with that previously given of interest:



This material is discussed in §12.4 of Nagel Saff and Snider. The method of Lyaponov as discussed in §12.5 is a way of generalizing this energy method to autonomous ODEs which are not direct reductions of Newton's equation. That is a very interesting topic, but we don't go too deep here. Let us conclude our brief study of qualitative methods with a discussion of **homogeneous constant coefficient linear systems**. The problem $\frac{d\vec{x}}{dt} = A\vec{x}$ we solved explicitly by the generalized e-vector method and we can make some general comments here without further work:

1. if all the e-values were both negative then the solutions will tend towards $(0, 0)$ as $t \to \infty$ due to the exponentials in the solution.

2. if any of the e-values were positive then the solutions will be unbounded as $t \to \infty$ since exponentials in the solution.

3. if the e-value was pure imaginary then the motion is bounded since the formulas are just sines and cosines which are bounded

4. if the e-value was complex with negative real part then the associated motion is stable and tends to $(0, 0)$ as the exponentials damp the sines and cosines in the $t \to \infty$ limit.

5. if the e-value was complex with positive real part then the associated motion is unstable and becomes unbounded as the exponentials blow-up in the limit $t \to \infty$.

See the table in §12.2 on page 779 of Nagel Saff and Snider for a really nice summary. Note however, my comments apply just as well to the $n = 2$ case as the $n = 22$ case. In short, the **spectrum** of the matrix A determines the stability of the solutions for $\frac{d\vec{x}}{dt} = A\vec{x}$. The spectrum is the list of the e-values for A. We could explicitly prove the claims I just made above, it ought not be too hard given all the previous calculations we've made to solve the homoegeneous constant coefficient case. What follows is far less trivial.

Theorem 5.1.3.

If the almost linear system $\frac{d\vec{x}}{dt} = A\vec{x} + \vec{f}(t, \vec{x})$ has a matrix A with e-values whose real parts are all negative then the zero solution of the almost linear system is asymptotically stable. However, if A has even one e-value with a positive real part then the zero solution is unstable.

This theorem is due to Poincare and Perron as is stated in section §12.7 page 824 of Nagel Saff and Snider. Here is a sketch of the idea behind the theorem:

1. when $\vec{x}$ is sufficiently small the term $\vec{f}(t, \vec{x})$ tends to zero hence the ODE is well-approximated by $\frac{d\vec{x}}{dt} = A\vec{x}$
2. close to the origin the problem is essentially the same as that we have already solved thus the e-values reveal the stability or instability of the origin.

In the pure imaginary case the theorem is silient because it is not generally known whether that pure cyclicity of the localization of the ODE will be maintained globally or if it will be spoiled into spiral-type solutions. Spirals can either go out or in and that is the trouble for the pure imaginary case.

This idea is just another appearance of the linearization concept from calculus. We can sometimes replace a complicated, globally nonlinear ODE, with a simple almost linear system. The advantage is the usual one; linear systems are easier to analyze.

In any event, to go deeper into these matters it would be wiser to think about manifolds and general coordinate change since we are being driven to think about such issues like it or not. Have no fear, your course ends here.

5.2 practice problems

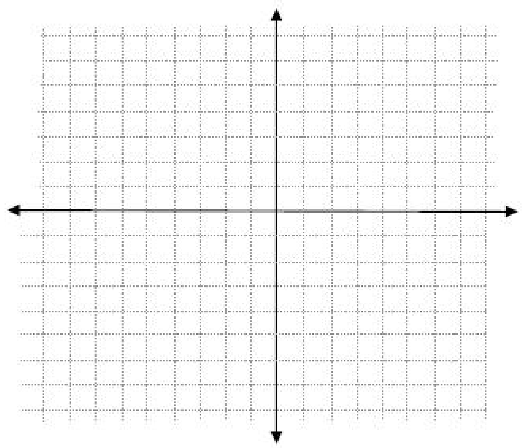
PP 378 Suppose a force $F(x) = 3x^4 + 16x^3 + 6x^2 - 72x$ is the net-force on some mass $m = 1$. Newton's Equation is $\ddot{x} = 3x^4 + 16x^3 + 6x^2 - 72x$.

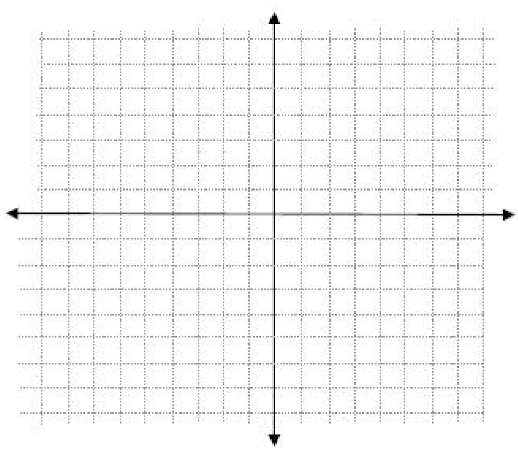
(1.) make the substitution $v = \dot{x}$ and write Newton's equation as a system in normal form for x and v.

(2.) find all three critical points for the system in (1.). (the potential should factor nicely)

(3.) plot the potential plane and phase plane juxtaposed vertically with the potential at the top and the phase plane at the base. Plot several trajectories and include arrows to indicate the direction of physically feasible solutions.

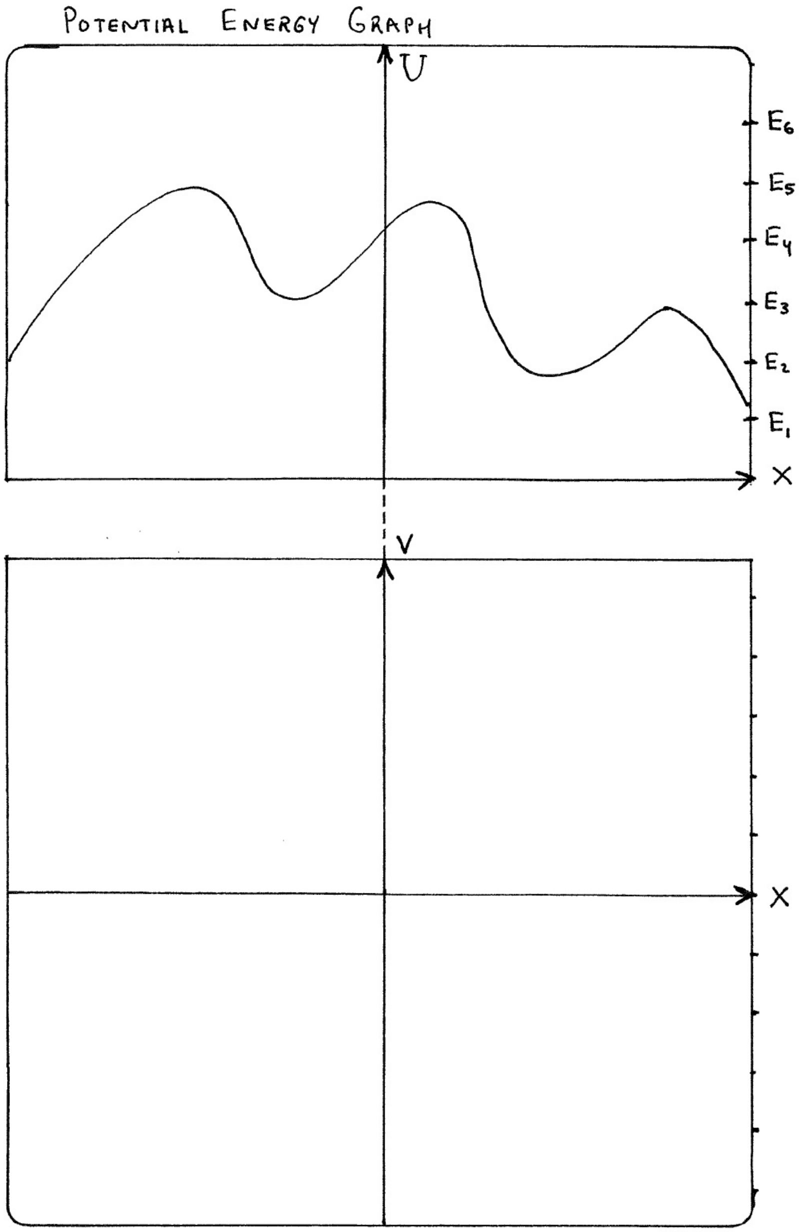
(4.) classify each critical point by examining your plot from (3.)



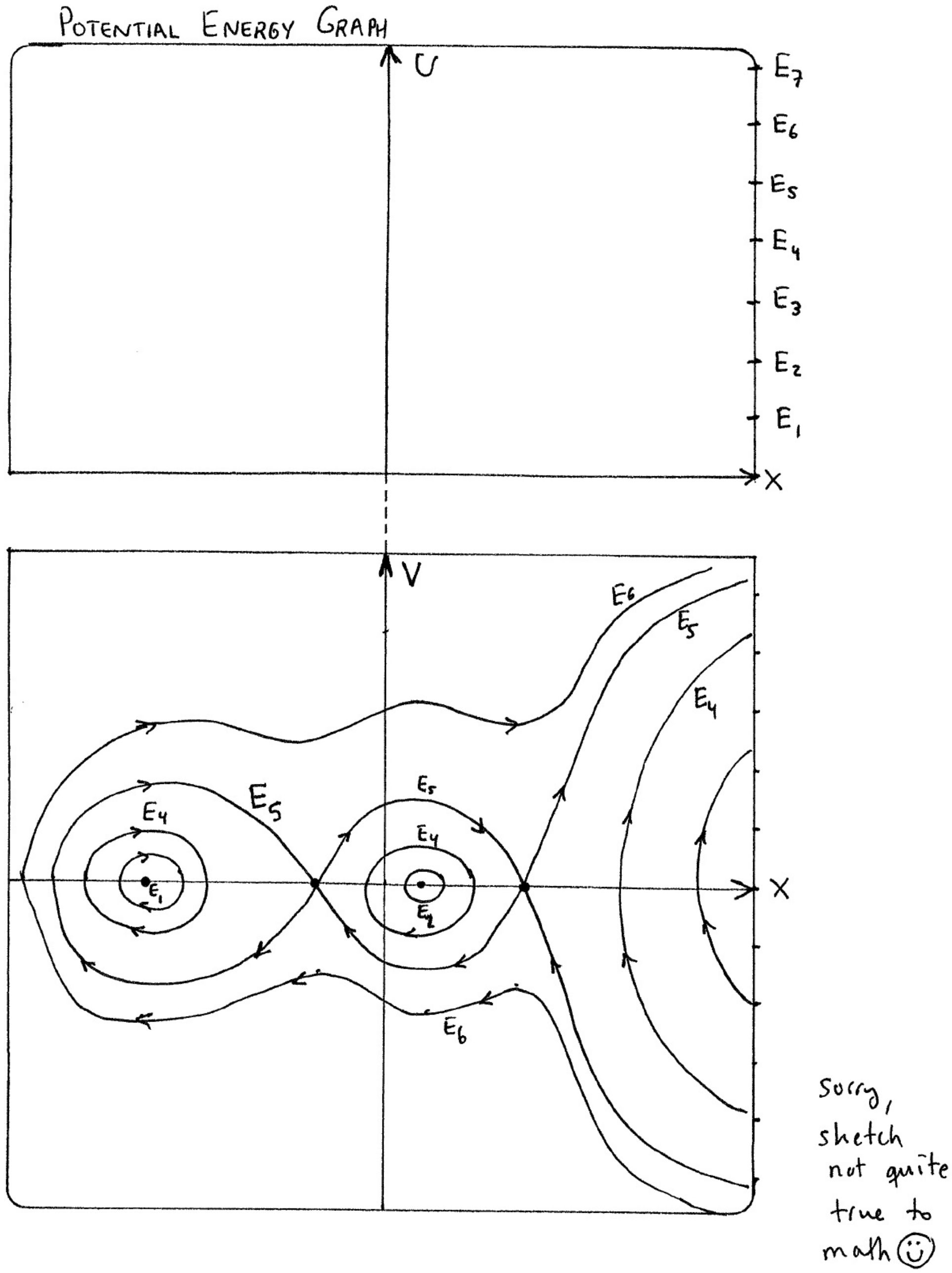


in this context the phase plane is also called the Poincare plane in honor of the mathematician who did much pioneering work in this realm of qualitative analysis. Incidentally, given any autonomous system $\frac{dx}{dt} = g(x, y)$ *and* $\frac{dy}{dt} = f(x, y)$ *we can study the timeless phase plane equation* $\frac{dy}{dx} = \frac{f}{g}$ *to indirectly analyze the solutions to the system. Solutions to the phase plane equation are the Cartesian level curves which are parametrized, with parameter* t*, by the solutions to the system*

PP 379 Plot the phase plane (or Poincare plot) given the potential energy plot below. For each energy $E_1, E_2, \ldots, E_6$ graph the corresponding trajectories below. Use a couple different colors so your work is easy to follow. Be neat. If no motion is possible then explain why.



PP 380 You are given a not so great phase plane (or Poincare plot) of the motion of a particle with various energies as listed. Plot the potential energy responsible for such motion.



Chapter 6

The Laplace transform technique

In our treatment of constant coefficients of differential equations, we discovered that we can translate the problem of calculus into a corresponding problem of algebra. Laplace transforms do something similar, however, Laplace transforms allow us to solve a wider class of problems. In particular, the Laplace transform will allow us to derive an elegant solution to problems with discontinuous forcing functions ($g(x)$ is the forcing function). In short, the method of Laplace transforms provides a powerful method to solve a wide class of ODE's which appear in common applications (especially electrical engineering, where t is time and s is frequency).
The motivation for the method of Laplace transforms is not the easiest thing to find in many texts on differential equations. However, there is actually a nice intuition for it. I think the best thing to do is to hear it from the master:

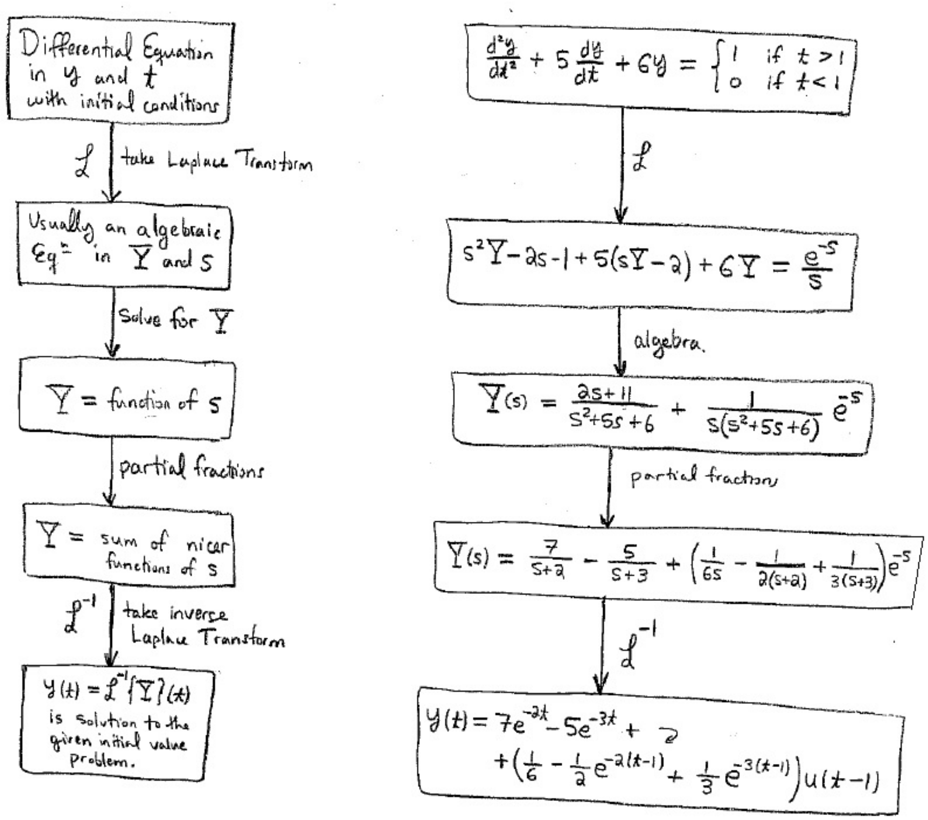
http://ocw.mit.edu/courses/mathematics/18-03-differential-equations-spring-2010/video-lectures/lecture-19-introduction-to-the-laplace-transform/

Check it out for yourself.

6.1 History and overview

We will spend the next few lectures learning how to take Laplace transforms and inverse Laplace transforms. The reason is primarily the following: We can solve differential equations with *discontinuous* forcing equations in a clean elegant manner.

Example 6.1.1. *This example gives you a look ahead to our end goal in this Chapter: using the Laplace transform to solve initial value problems for ordinary differential equations:*



The function $u(t-1) = \begin{cases} 1, & \text{if } t > 1 \\ 0, & \text{if } t < 1 \end{cases}$ is known as the "unit-step" or "Heaviside" function in honor of Oliver Heaviside who was one of the pioneers in this sort of Mathematics.

6.2 Definition and existence

Definition 6.2.1. *Laplace Transform*

Let $f(t)$ be a function on $[0, \infty)$. The **Laplace transform** of f is the function F defined by:

$$F(s) := \int_0^\infty e^{-st} f(t)\, dt := \mathcal{L}(f)(s)$$

F is defined on all real numbers s such that the intergral is finite.

Remark 6.2.2.

You should recall that $\int_0^\infty g(t)\, dt = \lim_{n\to\infty} \int_0^n g(t)\, dt$. In this course we use both of the following notations:

$$\int_0^\infty e^{-x}\, dx = \lim_{n\to\infty} \int_0^n e^{-x}\, dx = \lim_{n\to\infty} (-e^n + 1) = 1 \qquad \text{(explicitly)}$$

$$\int_0^\infty e^{-x}\, dx = -e^{-x}\Big|_0^\infty = e^{-\infty} + 1 = 1 \qquad \text{(implicitly)}$$

If I ask you to be explicit then follow the direction, you will likely to see the implicit version in applied courses. The implicit version is usually ok, but when something subtle arises it will confuses or disguise the issue. for example, what is $\infty e^{-\infty}$?

Theorem 6.2.3.

The Laplace transform $\mathcal{L}$ is a linear operator. In particular,

$$\mathcal{L}(f + g) = \mathcal{L}(f) + \mathcal{L}(g) \qquad \& \qquad \mathcal{L}(cf) = c\mathcal{L}(f).$$

Proof: Follows imediately from the linearity of integral. ☐

To be picky for a moment, you might ask which f and g are allowed in the above Theorem? Can we take the Laplace transform of any function? Skip ahead to Definition 6.2.9 to see a useful criteria for when a function f has a Laplace transform.

Example 6.2.4. Problem: Calculate the Laplace transform of the constant function $f(t) = 1$.

Solution: *calculate directly from the definition as follows:*

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt = \int_0^\infty e^{-st}\, dt = \frac{-1}{s} e^{-st}\Big|_0^\infty = \frac{1}{s}$$

for $s > 0$. *Thus* $\mathcal{L}(1)(s) = \frac{1}{s}$.

Example 6.2.5. *Find the Laplace transform of the function* $f(t) = e^{at}$.

Solution: *calculate directly from the definition as follows:*

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt = \int_0^\infty e^{(-s+a)t}\, dt = \frac{-1}{s-a} e^{(-s+a)t}\Big|_0^\infty = \frac{1}{s-a}$$

for $s > a$. *Thus* $\mathcal{L}(e^{at})(s) = \frac{1}{s-a}$.

The calculation of the example above is also valid if we replace a with $\alpha + i\beta$. Observe that

$$\frac{1}{s-\alpha-i\beta} = \frac{1}{s-\alpha-i\beta} \cdot \frac{s-\alpha+i\beta}{s-\alpha+i\beta} = \frac{s-\alpha+i\beta}{(s-\alpha)^2+\beta^2}$$

Example 6.2.6. Problem: Find the Laplace transform of the functions $f(t) = e^{\alpha t}\cos bt$ and $f(t) = e^{\alpha t}\sin bt$, where $\alpha, \beta \in \mathbb{R}$ with $\beta \neq 0$.

Solution: *the linearity of the Laplace transform applies in the complex case. We can reason:*

$$\mathcal{L}(e^{(\alpha+i\beta)t}) = \mathcal{L}(e^{\alpha t}\cos(\beta t) + ie^{\alpha t}\sin(\beta t) = \mathcal{L}(e^{\alpha t}\cos(\beta t)) + i\mathcal{L}(e^{\alpha t}\sin(\beta t)).$$

However, following the algebra above this example,

$$\mathcal{L}(e^{\alpha t}\cos(\beta t)) + i\mathcal{L}(e^{\alpha t}\sin(\beta t)) = \frac{s-\alpha+i\beta}{(s-\alpha)^2+\beta^2}.$$

We equate real and imaginary parts of the complex equation above to derive:

$$\mathcal{L}(e^{\alpha t}\cos(\beta t)) = \frac{s-\alpha}{(s-\alpha)^2+\beta^2} \qquad \& \qquad \mathcal{L}(e^{\alpha t}\sin(\beta t)) = \frac{\beta}{(s-\alpha)^2+\beta^2}.$$

Example 6.2.7. Problem: Find the Laplace transform of the function $f(t) = \begin{cases} 1, & \textbf{if } 0 \leq t < a \\ 0, & \textbf{if } t > a \end{cases}$.

Solution: *we proceed by direct calculation from the definition:*

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt = \int_0^a e^{-st} \cdot 0\, dt + \int_a^\infty e^{-st} \cdot 1\, dt = \frac{e^{-sa}}{s}.$$

The function $f(t) = H(t-a)$ is an example of the Heaviside or unit step function. Notice that it is discontinuos at $t = a$.

Remark 6.2.8.

You may recall that any continuous function is integrable. In fact, it is possible to integrate any function with finitely many jump-discontinuities. You just break up the integral into piecies, the value of the function at the discontinuities is irrelevant to integration.

One may inquire, for which functions on $[0, \infty)$ does the Laplace transform exist. Certainly piecewise continuity is a convenient assumption for the integral over a finite domain to exist. In addition we need $f(t)$ not to grow too fast or else the integral of the Laplace transform will diverge.

Definition 6.2.9.

A function $f(t)$ is said to be exponential order α iff there exists positive constanst T, M such that $|f(t)| \leq Me^{\alpha t}$ for all $t \geq T$.

This criteria will allow us to state when the Laplace transform of $f(t)$ exists, i.e. when the integral is finite.

Theorem 6.2.10.

If $f(t)$ is piecewise continuous on $[0, \infty)$ and of exponential of order α, then $\mathcal{L}(f)(s)$ exists for $s > \alpha$.

Proof: let f be piecewise continuous on $[0, \infty)$ with exponential order α. Then:

$$\int_0^\infty e^{-st} f(t)\, dt \le \int_0^\infty e^{-st}|f(t)|\, dt \le \int_0^T e^{-st}|f(t)|\, dt + \int_T^\infty e^{-st} M e^{\alpha t}\, dt \le C + Me^{-(s-\alpha)t}\Big|_\infty^T \le \infty$$

Thus $\mathcal{L}(f)(s)$ exists. $\square$

Examples of Laplace transformable functions:

(1.) e^{at} has exponential of order a.

(2.) $\sin t$ has exponential of order 0.

(3.) $\cos bt$ has exponential of order 0.

(4.) $e^{at} \cos bt$ is exponential of order a.

(5.) t^n is exponential of order 1 ($|t^n| < e^t$, for all $t > 1$)

We observe that all the functions which appear as fundamental solution of constant coefficient ODE's can be Laplace transformed. This is good as it is necessary if $\mathcal{L}$ is to handle common examples. The following table summarizes some popular Laplace transforms. I invite the interested reader to derive those which we have not already calculated in this section.

Known Laplace transforms: $\mathcal{L}(f)(s) := F(s)$

$f(t)$	$F(s)$	$\text{dom}(F)$
1	$\frac{1}{s}$	$s > a$
e^{at}	$\frac{1}{s-a}$	$s > a$
t^n	$\frac{n!}{s^{n+1}}$	$s > 0$
$\sin bt$	$\frac{b}{s^2+b^2}$	$s > 0$
$\cos bt$	$\frac{s}{s^2+b^2}$	$s > 0$
$e^{at}t^n$	$\frac{n!}{(s-a)^{n+1}}$	$s > a$
$e^{at} \sin bt$	$\frac{b}{(s-a)^2+b^2}$	$s > a$
$e^{at} \cos bt$	$\frac{s-a}{(s-a)^2+b^2}$	$s > a$

Examples are always helpful to assimilate the concepts here.

Example 6.2.11. Problem: Find the Laplace transform of the function $f(t) = t^2 - 3t - 2e^{-t} \sin 3t$**.**

Solution: *begin by using the linearity of the transform:*

$$\begin{aligned} F(s) &= \mathcal{L}(t^2)(s) + \mathcal{L}(-3t)(s) + \mathcal{L}(-2e^{-t} \sin 3t)(s) \\ &= \frac{2}{s^3} - \frac{3}{s^2} - \frac{6}{(s+1)^2 + 9}. \end{aligned}$$

Example 6.2.12. Problem: Find the Laplace transform of the function $f(t) = e^{-2t}\cos\sqrt{3}t - t^2e^{-2t}$.

Solution: *by linearity of the transform we calculate:*

$$F(s) = \mathcal{L}(e^{-2t}\cos\sqrt{3}t)(s) - \mathcal{L}(t^2e^{-2t})(s) \\ = \frac{s+2}{(s+2)^2+3} - \frac{2}{(s+2)^3}.$$

Taking the Laplace transform is not too hard. The trouble comes later when we try to go backwards.

Theorem 6.2.13. *Shift Theorem:*

> If the Laplace transform $\mathcal{L}(f)(s) = F(s)$ exists for $s > \alpha$, then for $s > \alpha + a$
>
> $$\mathcal{L}(e^{at}f(t))(s) = F(s-a)$$

Proof: is essentially just an algebra step:

$$\mathcal{L}(e^{at}f(t))(s) = \int_0^\infty e^{-st}e^{at}f(t)\,dt \\ = \int_0^\infty e^{-(s-a)t}f(t)\,dt \\ = F(s-a) \qquad \square$$

Example 6.2.14. *Observe,* $\mathcal{L}(e^{at}\sin bt)(s) = F(s-a) = \frac{b}{(s-a)^2+b^2}$ *is an example of the shift theorem in view of* $\mathcal{L}(\sin bt)(s) = \frac{b}{s^2+b^2} = F(s)$.

Example 6.2.15. *Observe* $\mathcal{L}(e^{at})(s) = F(s-a) = \frac{1}{s-a}$ *is an example of the shift theorem as* $\mathcal{L}(1)(s) = 1/s$.

Theorem 6.2.16. *Laplace Transform of Derivative:*

> Let f and f' be piecewise continuous with exponential order α, the for $s > \alpha$,
>
> $$\mathcal{L}(f')(s) = s\mathcal{L}(f)(s) - f(0)$$

Proof: is essentially by the product rule (a.k.a. as integration by parts):

$$\mathcal{L}(f')(s) = \lim_{n\to\infty}\int_0^n e^{-st}\frac{d}{dt}(f(t))\,dt \\ = \lim_{n\to\infty}\int_0^n \left(\frac{d}{dt}(e^{-st}f(t)) - \frac{d}{dt}(e^{-st})f(t)\right)dt \\ = \lim_{n\to\infty}\left(e^{-sn}f(n) - f(0) + s\int_0^n e^{-st}f(t)\,dt\right) \\ = -f(0) + s\mathcal{L}(f)(s)$$

Where we observed, for $t > \alpha$, we have $|f(t)| \leq Me^{\alpha t}$. Thus $|e^{-sn}f(n)| \leq e^{-sn}Me^{\alpha n} = Me^{n(\alpha-s)} \to 0$ as $n \to \infty$, provided that $s \geq \alpha$. $\square$

Similarly, we can derive the following theorem:

Theorem 6.2.17. *Laplace Transform of Derivative:*

Let $f, f', f'', ..., f^{(n-1)}$ be continuous and $f^{(n)}$ piecewise continuous, of exponential order α. Then for $s > \alpha$,

$$\mathcal{L}(f^{(n)})(s) = s^n \mathcal{L}(f)(s) - \sum_{i=1}^{n} s^{n-i} f^{(i-1)}(0)$$

Example 6.2.18. Problem: use the previous theorems to transform the constant coefficient differential equation $ay'' + by' + cy = 0$

$$\mathcal{L}(y'')(s) = s^2\mathcal{L}(y)(s) - sy(0) - y'(0)$$

$$\mathcal{L}(y')(s) = s\mathcal{L}(y)(s) - y(0)$$

$$\mathcal{L}(y)(s) = Y(s)$$

It is customary to use lower case y for $y(t)$, Y to denote the Laplace transform $\mathcal{L}(y)$. Taking the Laplace transform of $ay'' + by' + cy = 0$ yields:

$$a\left(s^2Y - sy(0) - y'(0)\right) + b\left(sY - y(0)\right) + cY = 0$$

or collecting terms with Y

$$(as^2 + bs + c)Y = asy(0) + ay'(0) + by(0).$$

Notice we have transformed a differential equation in t to an algebraic equation in s. Beautiful.

Example 6.2.19. Problem: Let $g(t) = \int_0^t f(u)\, du$**, calculate the Laplace transform of** g

Solution: *observe $g'(t) = f(t)$ by the Fundamental theorem of Calculus. Thus*

$$\mathcal{L}(f)(s) = \mathcal{L}(g')(s) = s\mathcal{L}(g)(s) - g(0) = sG(s)$$

Therefore,

$$\frac{1}{s}\mathcal{L}(f)(s) = \mathcal{L}\left(\int_0^t f(u)\, du\right)(s).$$

We just saw that integration in the t-domain prompts division by s of the transform of the integrand. On the other hand, $\mathcal{L}(dy/dt)(s) = sY - y(0)$ indicates differentiation in the t-domain prompts a multiplication by s in the s-domain. The theorem that follows shows there is also a relation between differentiation in the s-domain and multiplication by t in the t-domain:

Theorem 6.2.20.

Let $F(s) = \mathcal{L}(f)(s)$ and assume f is piecewise continous on $[0, \infty)$ and of exponential order α. Then for $s > \alpha$,

$$\mathcal{L}(t^n f(t))(s) = (-1)^n \frac{d^n F}{ds^n}(s).$$

Proof: Let us investigate the $n = 1$ case:

$$\begin{aligned}\frac{dF}{ds} &= \frac{d}{ds}\left(\int_0^\infty e^{-st} f(t)\, dt\right) \\ &= \int_0^\infty \frac{d}{ds}\left(e^{-st} f(t)\right) dt \\ &= \int_0^\infty -te^{-st} f(t)\, dt \\ &= (-1)\mathcal{L}(tf(t))(s)\end{aligned}$$

For the case $n = 2$,

$$\begin{aligned}\frac{d^2F}{ds^2} &= \frac{d}{ds}\left(\frac{dF}{ds}\right) \\ &= (-1)\frac{d}{ds}\int_0^\infty e^{-st} t f(t)\, dt \\ &= (-1)\int_0^\infty \frac{d}{ds}\left(e^{-st} t f(t)\right) dt \\ &= (-1)^2\int_0^\infty e^{-st} t^2 f(t)\, dt \\ &= (-1)^2\mathcal{L}(t^2 f(t))(s).\end{aligned}$$

The arguments for $n = 3, 4, \ldots$ are similar and could be made rigorous via an inductive argument. □

Summary of Laplace Transforms Theorems:

(1.) $\mathcal{L}(f + g) = \mathcal{L}(f) + \mathcal{L}(g)$
(2.) $\mathcal{L}(cf) = c\mathcal{L}(f)$
(3.) $\mathcal{L}(e^{at} f(t))(s) = \mathcal{L}(f)(s - a)$
(4.) $\mathcal{L}(f^{(n)})(s) = s^n\mathcal{L}(f)(s) - \sum_{i=1}^{n} s^{n-i} f^{(i-1)}(0)$
(5.) $\mathcal{L}(t^n f(t))(s) = (-1)^n F^{(n)}(s)$

Example 6.2.21. Problem: Let $f(t) = (1 + e^{-t})^2$ calculate $\mathcal{L}(f)(s)$

Solution: *observe* $f(t) = 1 + 2e^{-t} + e^{-2t}$ *hence:*

$$\begin{aligned}\mathcal{L}(f)(s) &= \mathcal{L}(1)(s) + \mathcal{L}(2e^{-t})(s) + \mathcal{L}(e^{-2t})(s) \\ &= \frac{1}{s} + \frac{2}{s+1} + \frac{1}{s+2}.\end{aligned}$$

Example 6.2.22. Problem: Let $f(t) = te^{2t}\cos 5t$. Calculate $\mathcal{L}(f)(s)$.

Solution: *utilize the theorem which says multilpication by* t *amounts to* $-\frac{d}{ds}$ *of the transform:*

$$\begin{aligned}\mathcal{L}(f)(s) &= -\frac{d}{ds}\left(\mathcal{L}(e^{2t}\cos 5t)(s)\right) \\ &= -\frac{d}{ds}\left(\frac{s-2}{(s-2)^2+25}\right) \\ &= \frac{(s-2)^2-25}{((s-2)^2+25)^2}\end{aligned}$$

Example 6.2.23. Problem: Let $f(t) = t\sin^2 t$. Calculate $\mathcal{L}(f)(s)$.

Solution: *begin by applying the appropriate theorem, then recall a trigonometric identity:*

$$\begin{aligned}
\mathcal{L}(f)(s) &= -\frac{d}{ds}\left(\mathcal{L}(\sin^2 t)(s)\right) \\
&= -\frac{d}{ds}\left(\mathcal{L}\left(\frac{1-\cos 2t}{2}\right)(s)\right) \\
&= -\frac{1}{2}\frac{d}{ds}\left(\frac{1}{s} - \frac{s}{s^2+4}\right) \\
&= \frac{1}{2}\left(\frac{1}{s^2} + \frac{4-s^2}{(s^2+4)^2}\right)
\end{aligned}$$

Example 6.2.24. Problem: Calculate Laplace transform of $f(t) = \sin t \cos 2t$.

Solution: *I use the technique of imaginary exponentials to derive the necessary trigonometric identity:*

$$\begin{aligned}
\sin t \cos 2t &= \frac{1}{2i}\left(e^{it} - e^{-it}\right)\frac{1}{2}\left(e^{2it} + e^{-2it}\right) \\
&= \frac{1}{4i}\left(e^{3it} - e^{-3it} + e^{-it} - e^{it}\right) \\
&= \frac{1}{2}\sin 3t - \frac{1}{2}\sin t.
\end{aligned}$$

Thus, in view of the algebra above,

$$\mathcal{L}(f)(s) = \frac{3}{2(s^2+9)} - \frac{1}{2(s^2+1)}.$$

Example 6.2.25. Problem: Calculate Laplace transform of $f(t) = te^{-t} - 3$.

Solution: *use linearity and the known transforms.*

$$\mathcal{L}(te^{-t} - 3)(s) = \mathcal{L}(te^{-t})(s) + \mathcal{L}(-3)(s) = \frac{1}{(s+1)^2} - \frac{3}{s}.$$

Example 6.2.26. Problem: Calculate Laplace transform of $f(t) = 13e^{2t}\sin(t+\pi)$.

Solution: *I use the trigonometric identity* $\sin(a+b) = \sin a \cos b + \cos a \sin b$ *to make the calculation below:*

$$\mathcal{L}(13e^{2t}\sin(t+\pi))(s) = -\mathcal{L}(13e^{2t}\sin t)(s) = -\frac{13}{(s-2)^2+1}.$$

Example 6.2.27. Problem: Calculate Laplace transform of $f(t) = te^{-2t}\sin 3t$.

Solution: *we use the multiplication by t gives $-d/ds$ of transform theorem:*

$$\begin{aligned}\mathcal{L}(te^{-2t}\sin 3t)(s) &= -\frac{d}{ds}\left(\mathcal{L}(e^{-2t}\sin 3t)(s)\right) \\ &= -\frac{d}{ds}\left[\frac{3}{(s+2)^2+9}\right] \\ &= \frac{6(s+2)}{\left[(s+2)^2+9\right]^2}.\end{aligned}$$

Example 6.2.28. Problem: Calculate Laplace transform of $f(t) = \sin t \cos^2 t$.

Solution: *begin by deriving the needed trigonometric identity:*

$$\begin{aligned}\sin t \cos^2 t &= \frac{e^{it}-e^{-it}}{2i}\left(\frac{e^{it}+e^{-it}}{2}\right)^2 \\ &= \frac{1}{8i}(e^{it}-e^{-it})(e^{2it}+2+e^{-2it}) \\ &= \frac{1}{8i}(e^{3it}+2e^{it}+e^{-it}-e^{it}-2e^{-it}-e^{-3it}) \\ &= \frac{1}{4}\frac{1}{2i}(e^{3it}-e^{-3it})+\frac{1}{4}\frac{1}{2i}(e^{it}-e^{-it}) \\ &= \frac{1}{4}\sin 3t+\frac{1}{4}\sin t\end{aligned}$$

All that remains is to apply the known transforms:

$$\mathcal{L}(\sin t\cos^2 t)(s) = \frac{1}{4}\left(\frac{3}{s^2+9}+\frac{1}{s^2+1}\right).$$

6.3 The inverse transform

Definition 6.3.1.

Given $F(s)$, if there is a function $f(t)$ on $[0,\infty)$ such that $\mathcal{L}(f) = F$ then we say f is the inverse Laplace transform of F and denote $f = \mathcal{L}^{-1}(F)$

Remark 6.3.2.

There are many possible choices for f given some particular F. This is due to the fact that $\int_0^\infty e^{-st}f_1(t)\,dt = \int_0^\infty e^{-st}f_2(t)\,dt$ provided that f_1 and f_2 disagree only at a few points. The result of the inverse transform is unique if we require f to be continuous. To those who know measure theory, this means that f is unique up to an equivalence relation on $l^\infty([0,\infty))$. The relation is define as following: $f \sim g$ iff $f(t) = g(t)$ almost everywhere, i.e. the Lebesgue measure of the set of points where f and g differ is 0. This is a subtle point and I've already said too much here.

Example 6.3.3. Problem: *calculate the inverse transform of $F(s) = \frac{2}{s^3}$.*

Solution: *observe $\mathcal{L}(t^2)(s) = \frac{2}{s^3}$ thus $\mathcal{L}^{-1}(2/s^3)(t) = t^2$.*

Example 6.3.4. **Problem:** *calculate the inverse transform of* $F(s) = \frac{3}{s^2+9}$.

Solution: *recall* $\mathcal{L}(\sin 3t)(s) = \frac{3}{s^2+9}$ *thus* $\mathcal{L}^{-1}\left(\frac{3}{s^2+9}\right)(t) = \sin 3t$.

Example 6.3.5. **Problem:** *calculate the inverse transform of* $F(s) = \frac{s-1}{s^2-2s+5}$.

Solution: *completing the square reveals much:*

$$\mathcal{L}^{-1}\left(\frac{s-1}{s^2-2s+5}\right)(t) = \mathcal{L}^{-1}\left(\frac{s-1}{(s-1)^2+4}\right)(t) = e^t\cos 2t.$$

Reminder: To complete the square, we simply want to rewrite a quadratic form x^2+bx+c into $(x-h)^2+k$. To do this we just take $\frac{1}{2}$ of coefficient of x and then $(x+\frac{b}{2})^2 = x^2+bx+\frac{b^2}{4}$ so we then have to subtract $\frac{b^2}{4}$ to be fair,

$$x^2+bx+c = \left(x+\frac{b}{2}\right)^2 - \frac{b^2}{4} + c.$$

It's easier to understand for specific examples,

$$x^2+2x+5 = (x+1)^2 - 1 + 5 = (x+1)^2 + 4$$
$$x^2+6x+5 = (x+3)^2 - 9 + 5 = (x+3)^2 - 4$$

In practice I just make sure that the LHS and RHS are equal, you don't need to remember some algorithm if you understand the steps.

Theorem 6.3.6.

Inverse Laplace Transform is linear provided we choose continuous f

$$\mathcal{L}^{-1}(F+G) = \mathcal{L}^{-1}(F) + \mathcal{L}^{-1}(G),$$

$$\mathcal{L}^{-1}(cF) = c\mathcal{L}^{-1}(F).$$

Proof: It follows from linearity of $\mathcal{L}$ that

$$\mathcal{L}(\mathcal{L}^{-1}(F) + \mathcal{L}^{-1}(G)) = \mathcal{L}(\mathcal{L}^{-1}(F)) + \mathcal{L}(\mathcal{L}^{-1}(G)) = F + G$$

Then $\mathcal{L}^{-1}(F+G) = \mathcal{L}^{-1}(F) + \mathcal{L}^{-1}(G)$. Similar argument for homogeneous property. □

If we drop the stipulation of continuity then linearity holds *almost everywhere* for the inverse transforms.

Example 6.3.7. **Problem:** *find* $f(t) = \mathcal{L}^{-1}(F)(t)$ *for* $F(s) = \frac{3s+2}{s^2+2s+10}$.

Solution: *the algebra below anticipates the inverse transform:*

$$\begin{aligned}\frac{3s+2}{s^2+2s+10} &= \frac{3s+2}{(s+1)^2+9} \\ &= \frac{3(s+1)}{(s+1)^2+9} + \frac{-3+2}{(s+1)^2+9} \\ &= 3\frac{s+1}{(s+1)^2+3^2} - \frac{1}{3}\frac{3}{(s+1)^2+3^2}\end{aligned}$$

In view of the algebra above it should be obvious that

$$\mathcal{L}^{-1}(F)(t) = 3e^{-t}\cos 3t - \frac{1}{3}e^{-t}\sin 3t = f(t).$$

Example 6.3.8. Problem: *consider* $F(s) = \frac{s}{s^2+5s+6}$. *Find* $\mathcal{L}^{-1}(F)(t) = f(t)$.

Solution: *Note the following partial fractions decomposition:*

$$\frac{s}{s^2+5s+6} = \frac{s}{(s+2)(s+3)} = \frac{3}{s+3} - \frac{2}{s+2}.$$

Therefore, we deduce

$$\begin{aligned}\mathcal{L}^{-1}(F)(t) &= \mathcal{L}^{-1}\left(\frac{3}{s+3} - \frac{2}{s+2}\right)(t) \\ &= 3\mathcal{L}^{-1}\left(\frac{1}{s+3}\right)(t) - 2\mathcal{L}^{-1}\left(\frac{1}{s+2}\right)(t) \\ &= 3e^{-3t} - 2e^{-2t}.\end{aligned}$$

Partial Fractions: We have discussed how polynomials split into linear and irreducible quadratic factors. This means if we have a rational function which is $\frac{p(s)}{g(s)}$ then $p(s)$ and $g(s)$ will factor, we assume $deg(p) < deg(g)$ for convenience (otherwise we would do long division). In short, partial fractions says you can split up a rational function into a sum of basic rational functions. For basic rational functions we can readily see how to take the inverse transform. Partial fractions involves a number of cases as you may read in most basic DEqns texts. I do not attempt generality here, I just wish for you to realize it is nothing more than undoing making a common denominator. I will leave you with a few examples,

$$\frac{s^3-3}{(s+1)^3(s^2+1)} = \frac{A}{s+1} + \frac{B}{(s+1)^2} + \frac{C}{(s+1)^3} + \frac{Ds+E}{s^2+1}$$

$$\frac{s+3}{s^2(s-2)(s^2+3)^2} = \frac{A}{s} + \frac{B}{s^2} + \frac{C}{s-2} + \frac{Ds+E}{s^2+3} + \frac{Fs+G}{(s^2+3)^2}$$

It is a simple, but tedious matter to calculate the constants $A, B, C, .., G$ in the above. Notice on the RHS almost all the terms are easily inverse transformed. The last term in the second equation is subtle, just as in the corresponding algebra problem we saw in the study of integration in calculus II.

Remark 6.3.9.

It is crucial to understand the difference between (s^2+1) and $(s+1)^2$. Note that the inverse transforms of $\left(\frac{1}{s^2+1}\right)$ and $\left(\frac{1}{(s+1)^2}\right)$ are quite different.

Example 6.3.10. Problem: *find the inverse Laplace transform of* $F(s) = \frac{s+1}{s^2-2s+5}$.

Solution: *the observation* $1 = -1 + 2$ *is especially useful below:*

$$F(s) = \frac{s+1}{s^2-2s+5} = \frac{s+1}{(s-1)^2+4} = \frac{(s-1)+2}{(s-1)^2+4}.$$

Therefore, $\mathcal{L}^{-1}(F)(t) = e^t\cos 2t + e^t\sin 2t$.

Example 6.3.11. Problem: *find the inverse Laplace transform of* $F(s) = \frac{7s^3-2s^2-3s+6}{s^3(s-2)}$.

Solution: *we propose the partial-fractal-decomposition below:*

$$F(s) = \frac{7s^3 - 2s^2 - 3s + 6}{s^3(s-2)} = \frac{A}{s} + \frac{B}{s^2} + \frac{C}{s^3} + \frac{D}{s-2}$$

multiply by the denominator $s^3(s-2)$ *to obtain:*

$$7s^3 - 2s^2 - 3s + 6 = As^2(s-2) + Bs(s-2) + C(s-2) + Ds^3.$$

There are a number of methods to solve such equations. As this must hold for all s *it is convienent to plug in a few simple integers:*

- **(i.)** $s = 0:$ *We have* $6 = -2C$. *Therefore* $C = -3$,
- **(ii.)** $s = 1:$ *We have* $8 = -A - B + 3 + D$,
- **(iii.)** $s = 2:$ *We have* $56 - 8 - 6 + 6 = 8D$. *Therefore* $D = 6$,
- **(iv.)** $s = -1:$ *We have* $-3A + 3B + 3 = 0$, *which implies* $A = B + 1$.

Hence (ii.) reduces to $8 = -B - 1 - B + 3 + 6$. *Therefore* $B = 0$ *and* $A = 1$. *In summary,*

$$F(s) = \frac{1}{s} - \frac{3}{s^3} + \frac{6}{s-2} \qquad \Rightarrow \qquad \mathcal{L}^{-1}(F)(t) = 1 - \frac{3t^2}{2} + 6e^{2t}.$$

Example 6.3.12. Problem: *find the inverse Laplace transform of* $F(s) = e^{-s}\frac{4s+2}{s(s+1)}$.

Solution: *Let us set* $F(s) = e^{-s}G(s)$. *Note that*

$$G(s) = \frac{4s+2}{s(s+1)} = \frac{2}{s} + \frac{2}{s+1}$$

Thus $g(t) = 2 + 2e^{-t}$. *Therefore,*

$$\begin{aligned} \mathcal{L}^{-1}(G(s)e^{-s})(t) &= g(t-1)u(t-1) \\ &= (2 + 2e^{-t+1})u(t-1) \\ &= 2(1 + e^{-t+1})u(t-1). \end{aligned}$$

6.3.1 how to solve an ODE via the method of Laplace transforms

Example 6.3.13. Problem: *solve* $y'' - 2y' + 5y = 0$ *with* $y(0) = 2$ *and* $y'(0) = 12$.

Solution: *Take the Laplace transform of the given initial value problem,*

$$s^2Y - sy(0) - y'(0) - 2(sY - y(0)) + 5Y = 0$$

We solve for Y,

$$(s^2 - 2s + 5)Y = 2s + 8 \qquad \Rightarrow \qquad Y(s) = \frac{2s+8}{s^2 - 2s + 5}.$$

We wish to find $y(t) = \mathcal{L}^{-1}(Y)(t)$. *To do that we need to find if* $s^2 - 2s + 5$ *will factor, note* $b^2 - 4ac = 4 - 20 = -16 < 0$. *Thus it is an irreducible quadratic. We will complete the square and break into* sin *and* cos *pieces.*

$$\frac{2s+8}{s^2-2s+5} = \frac{2s+8}{(s-1)^2+4} = \frac{2(s-1)}{(s-1)^2+2^2} + \frac{2}{2}\frac{8+2}{(s-1)^2+2^2}$$

Thus,

$$y(t) = 2e^t \cos 2t + 5e^t \sin 2t.$$

Example 6.3.14. Problem: *solve the repeated root problem* $y'' + 4y' + 4y = 0$ *subject to the intial conditions* $y(0) = 1$ *and* $y'(0) = 1$.

Solution: *the Laplace transform yields,*

$$s^2 Y - sy(0) - y'(0) + 4(sY - y(0)) + 4Y = 0$$

We solve for Y *as usual,*

$$(s^2 + 4s + 4)Y = s + 5 \qquad \Rightarrow \qquad Y = \frac{s+5}{s^2+4s+4} = \frac{1}{s+2} + \frac{3}{(s+2)^2}$$

Therefore,

$$y(t) = \mathcal{L}^{-1}(Y)(t) = \mathcal{L}^{-1}\left(\frac{1}{s+2}\right)(t) + \mathcal{L}^{-1}\left(\frac{3}{(s+2)^2}\right)(t)$$

and we conclude $y(t) = e^{-2t} + 3te^{-2t}$.

Remark 6.3.15.

The method of Laplace transform has derived the curious te^{-t} term. Before we just pulled it out of thin-air and argued that it worked. In defense of our earlier methods, the Laplace machine is not that intuitive either. At least we have one derivation now. Another route to explain the "t" in the double root solution is to use "reduction of order". There is also a pretty derivation based on the matrix exponential and generalized eigenvectors which we should discusss in our solution of systems of ODEs.

Example 6.3.16. Problem: $y'' + y = 2e^t$, *where* $y(0) = 1$ *and* $y'(0) = 2$.

Solution: *Take the Laplace transform,*

$$s^2 Y - s - 2 + Y = \frac{2}{s-1} \qquad \Rightarrow \qquad (s^2+1)Y = s + 2 + \frac{2}{s-1}$$

Algebra yields

$$Y = \frac{s+2}{s^2+1} + \frac{2}{(s^2+1)(s-1)} = \frac{s+2}{s^2+1} + \frac{-s-1}{s^2+1} + \frac{1}{s-1} = \frac{1}{s^2+1} + \frac{1}{s-1}$$

Taking $\mathcal{L}^{-1}$ *yields the solution:*

$$y = \sin t + e^t.$$

The standard Laplace technique is set-up for initial values at $t = 0$. However, the next example illustrates that is just a convenience of our exposition thus far.

Example 6.3.17. Problem: *solve* $w'' - 2w' + 5w = -8e^{\pi - t}$*, given* $w(\pi) = 2$ *and* $w'(\pi) = 12$.

Solution: *We need conditions at* $t = 0$ *so to remedy being given them at* π*. We define* y *as follows:*

$$y(t) = w(t + \pi) \Rightarrow y(0) = w(\pi) = 2$$

And,

$$y'(0) = w'(\pi) = 12$$

Then,

$$w''(t + \pi) - 2w'(t + \pi) + 5w(t + \pi) = -8e^{\pi - (t+\pi)} = -8e^{-t}$$

Thus,

$$y'' - 2y' + 5y = -8e^{-t}$$

Taking the Laplace transform yields,

$$(s^2 - 2s + 5)Y - 2s - 12 - 2(-2) = \frac{-8}{s+1}$$

$$\Rightarrow Y = \left(8 + 2s - \frac{8}{s+1}\right)\frac{1}{s^2 - 2s + 5} = \frac{3(s-1) + 2(4)}{(s-1)^2 + 2^2} - \frac{1}{s+1}$$

$$\Rightarrow \mathcal{L}^{-1}(Y)(t) = 3e^t \cos 2t + 4e^t \sin 2t - e^{-t} = y(t)$$

Then returning to our original problem. Note $w(t) = y(t - \pi)$*, thus:*

$$\begin{aligned} w(t) &= 3e^{t-\pi}\cos(2(t - \pi)) + 4e^{t-\pi}\sin(2(t - \pi)) - e^{\pi - t} \\ &= 3e^{t-\pi}\cos(2t - 2\pi) + 4e^{t-\pi}\sin(2t - 2\pi) - e^{\pi - t}. \end{aligned}$$

We find $w(t) = 3e^{t-\pi}\cos 2t + 4e^{t-\pi}\sin 2t - e^{\pi - t}$.

Remark 6.3.18.

This example is important in that it shows us how to use Laplace transforms to treat problems where the data is given at any time not just zero as the formalism is set-up for

Example 6.3.19. Problem: *suppose* $w(1) = 1$ *and* $w'(1) = 0$*. Solve:* $w'' - w = \sin(t - 1)$.

Solution: *Using Laplace transforms. Introduce* $y(t) = w(t + 1)$ *so that* $y(0) = w(1)$*. Also noteice that* $y'(t) = \frac{dw}{dt}\frac{d}{dt}(t + 1) = w'(t + 1)$ *and* $y''(t) = w''(t + 1)$*. Consider the differential equation at* $t + 1$*,*

$$w''(t + 1) - w(t + 1) = \sin(t + 1 - 1)$$

Thus, $y'' - y = \sin t$*, and* $y(0) = 1$*,* $y'(0) = 0$*. Now we can use the standard Laplace theory on* y*,*

$$s^2 Y - s - Y = \frac{1}{s^2 + 1}$$

$$\Rightarrow Y = \frac{1}{s^2 - 1}\left(s + \frac{1}{s^2 + 1}\right) = \frac{s(s^2 + 1) + 1}{(s+1)(s-1)(s^2+1)} = \frac{1}{4}\frac{1}{s+1} + \frac{3}{4}\frac{1}{s-1} - \frac{1}{2}\frac{1}{s^2+1}$$

Thus, $y(t) = \frac{1}{4}e^{-t} + \frac{3}{4}e^{t} - \frac{1}{2}\sin(t)$. *Therefore,*

$$w(t) = \frac{1}{4}e^{-t+1} + \frac{3}{4}e^{t-1} - \frac{1}{2}\sin(t-1)$$

There are other perhaps simpler ways to express our final answer, but this will suffice.

Example 6.3.20. Problem: *solve:* $ty'' - ty' + y = 2$, *where* $y(0) = 2$ *and* $y'(0) = -1$.

Solution: *We have:*

$$\mathcal{L}(ty'') = -\frac{d}{ds}\mathcal{L}(y'') = -\frac{d}{ds}(s^2Y - 2s + 1) = -2sY - s^2Y' + 2$$

And

$$\mathcal{L}(ty') = -\frac{d}{ds}\left(\mathcal{L}(y')\right) = -\frac{d}{ds}(sY - 2) = -Y - sY'$$

Thus, $\mathcal{L}(ty'' - ty' + y) = \mathcal{L}(2)$ *yields:*

$$\frac{dY}{ds}(s - s^2) + Y(2 - 2s) = \frac{2}{s} - 2$$

Hence,

$$-\frac{dY}{ds}s(s-1) - 2(s-1)Y = 2\left(\frac{1}{s} - 1\right) \quad \Rightarrow \quad s\frac{dY}{ds} + 2Y = \frac{2}{s}.$$

One more algebra step brings us to standard form:

$$\frac{dY}{ds} + \frac{2}{s}Y = \frac{2}{s^2}$$

Therefore, by the integrating factor technique:

$$Y = \frac{1}{s^2}\int s^2\frac{2}{s^2}\,ds = \frac{2}{s} + \frac{c_1}{s^2}$$

which implies

$$Y(s) = \frac{2}{s} + \frac{c_1}{s^2} \quad \Rightarrow \quad y(t) = 2 + c_1t.$$

Since $y'(0) = -1$, $c_1 = -1$. *As a result, we find the solution* $y(t) = 2 - t$.

There is probably an easier way to solve the problem above. I include it here merely to show you that there are problems where calculus must be done in the s-domain.

6.4 Discontinuous functions

One main motivation for including Laplace transforms in your education is that it allows us to treat problems with piecewise continuous forcing terms in a systematic fashion. Without this technique, you have to solve the problem in each piece then somehow glue them together.

Definition 6.4.1.

The unit-step function $u(t)$ is defined by

$$u(t) = \begin{cases} 0, & \text{if } t < 0 \\ 1, & \text{if } t > 0 \end{cases}$$

Often it will be convenient to use $u(t-a) = \begin{cases} 0, & \text{if } t < a \\ 1, & \text{if } t > a \end{cases}$. This allows us to switch functions on or off for particular ranges of t.

Example 6.4.2. *The casewise-defined function* $g(t) = \begin{cases} 0, & \textit{if } t < 0 \\ \cos t, & \textit{if } 0 < t < 1 \\ \sin t, & \textit{if } 1 < t < \pi \\ t^2, & \textit{if } \pi < t \end{cases}$ *can be written elegantly as:*

$$g(t) = (u(t) - u(t-1))\cos t + (u(t-1) - u(t-\pi))\sin t + u(t-\pi)t^2.$$

It is not hard to see why this function is useful to a myriad of applications, anywhwere you have a switch the unit-step functions provides an idealized model of that.

Proposition 6.4.3.

$\mathcal{L}(u(t-a))(s) = \frac{1}{s}e^{-as}$

Proof: We have

$$\mathcal{L}(u(t-a))(s) = \int_0^\infty e^{-st}u(t-a)\,dt = \int_a^\infty e^{-st}\,dt = \frac{-1}{s}e^{-st}\bigg|_a^\infty = \frac{1}{s}e^{-as} \quad \square.$$

Theorem 6.4.4.

Let $F(s) = \mathcal{L}(f)(s)$ for $s > \alpha \geq 0$. If $a > 0$, then

$$\mathcal{L}(f(t-a)u(t-a))(s) = e^{-as}F(s)$$

And conversely,

$$\mathcal{L}^{-1}(e^{-as}F(s))(t) = f(t-a)u(t-a)$$

Proof: We calculate from the definition,

$$\mathcal{L}(f(t-a)u(t-a))(s) = \int_0^\infty e^{-st}f(t-a)u(t-a)\,dt = \int_a^\infty e^{-st}f(t-a)\,dt =$$

$$= \int_0^\infty e^{-s(u+a)}f(u)\,du = e^{-sa}\mathcal{L}(f)(s) = e^{-as}F(s) \qquad \square.$$

Corollary 6.4.5.

$\mathcal{L}(g(t)u(t-a))(s) = e^{-as}\mathcal{L}(g(t+a))(s)$

Proof: Let $h(t-a) := g(t)$. The corollary is immediately follows from Theorem 6.4.4. $\square$

Example 6.4.6. *Simply apply the previous corollary to obtain,*

$$\mathcal{L}(t^2u(t-1))(s) = e^{-s}\mathcal{L}((t+1)^2)(s) = e^{-s}\mathcal{L}(t^2+2t+1)(s) =$$
$$= e^{-s}\left(\frac{2}{s^3}+\frac{2}{s^2}+\frac{1}{s}\right).$$

Example 6.4.7. Problem: *find* $\mathcal{L}^{-1}\left(\frac{1}{s^2}e^{-2s}\right)$.

Solution: *we use Theorem 6.4.4*

$$\mathcal{L}^{-1}\left(\frac{1}{s^2}e^{-2s}\right) = \mathcal{L}^{-1}\left(\frac{1}{s^2}\right)(t-2)u(t-2) = (t-2)u(t-2).$$

Remark 6.4.8.

If we find exponential factors in the s-domain, that suggests we will encounter unit-step functions upon taking $\mathcal{L}^{-1}$ to get back to the t-domain.

Example 6.4.9. Problem: Calculate Laplace transform of $f(t) = \begin{cases} e^t, & \textbf{if } 0 \le t \le 2 \\ t, & \textbf{if } t > 2 \end{cases}$**.**

Solution: *we begin by expressing the function in terms of unit-step functions:*

$$f(t) = e^t[u(t)-u(t-2)] + tu(t-2) = e^tu(t) + (t-e^t)u(t-2)$$

The transforms below follow, notice how the $u(t-2)$ *term prompts the appearance of the* $t+2$ *arguments:*

$$\begin{aligned}
\mathcal{L}(f)(s) &= \mathcal{L}(e^tu(t))(s) + \mathcal{L}((t-e^t)u(t-2))(s) \\
&= \mathcal{L}(e^t)(s) + e^{-2s}\mathcal{L}(t+2-e^{t+2})(s) \\
&= \mathcal{L}(e^t)(s) + e^{-2s}\mathcal{L}(t+2-e^2e^t)(s) \\
&= \frac{1}{s-1} + e^{-2s}\left(\frac{1}{s^2}+\frac{2}{s}-\frac{e^2}{s-1}\right).
\end{aligned}$$

Example 6.4.10. Problem: Calculate Laplace transform of $f(t) = \begin{cases} 1, & \textbf{if } \ 0 \le t \le 1 \\ t^2, & \textbf{if } 1 < t \le 2 \\ \sin t, & \textbf{if } t > 2 \end{cases}$**.**

Solution: *we begin by converting the formula to unit-step format:*

$$\begin{aligned}
f(t) &= (u(t)-u(t-1)) + (u(t-1)-u(t-2))t^2 + u(t-2)\sin t \\
&= u(t) + (t^2-1)u(t-1) + (\sin t - t^2)u(t-2)
\end{aligned}$$

We have: noting that $(t+2)^2 = t^2+4t+4$ *whereas* $(t+1)^2 = t^2+2t+1$,

$$\mathcal{L}(u(t))(s) = \frac{1}{s}$$
$$\mathcal{L}((t^2-1)u(t-1))(s) = e^{-s}\left(\frac{2}{s^3}+\frac{2}{s^2}\right)$$
$$\mathcal{L}((\sin t - t^2)u(t-2))(s) = e^{-2s}\left(\frac{\cos 2}{s^2+1}+\sin 2\frac{s}{s^2+1}-\frac{2}{s^3}-\frac{4}{s^2}-\frac{4}{s}\right)$$

Linearity combines the three results above to yield $\mathcal{L}(f)(s)$

$$\mathcal{L}(f)(s) = \frac{1}{s} + e^{-s}\left(\frac{2}{s^3} + \frac{2}{s^2}\right) + e^{-2s}\left(\frac{\cos 2}{s^2+1} + \sin 2\frac{s}{s^2+1} - \frac{2}{s^3} - \frac{4}{s^2} - \frac{4}{s}\right).$$

6.5 further Laplace transforms

I don't always have time to cover everything in this Section. Here I study periodic functions, the Gamma function and the extension of Laplace techniques to series. If you are truly curious then I would strongly advise you don't stop with this humble treatment. There is much more to learn[1]

Definition 6.5.1.

A function f is said to be periodic on its domain with period T if $f(t) = f(t+T)$ for all t.

Example 6.5.2. *note items (4.) and (5.) are extended periodically past the given window of the function.*

(1.) $f_1(t) = \sin t$ *has* $T = 2\pi$

(2.) $f_2(t) = \sin kt$ *has* $T = \frac{2\pi}{k}$

(3.) $f_3(t) = \tan t$ *has* $T = \pi$

(4.) $g(t) = \begin{cases} 1, & \text{if } 0 < t < 1 \\ 0, & \text{if } 1 < t < 2 \end{cases}$ *with* $T = 2$

(5.) $g(t) = \begin{cases} t, & \text{if } 0 < t < 1 \\ 2-t, & \text{if } 1 < t < 2 \end{cases}$ *with* $T = 2$

Definition 6.5.3.

For f with $[0,T] \in dom(f)$ with f periodic with period T, we define the "windowed version" of f

$$f_T(t) = \begin{cases} f(t), & \text{if } 0 < t < T \\ 0, & \text{otherwise} \end{cases}$$

The Laplace transform of the windowed version of a periodic function f with period T is similarly denoted

$$F_T(t) = \int_0^\infty e^{-st} f_T(t)\, dt = \int_0^T e^{-st} f(t)\, dt.$$

Theorem 6.5.4.

If f has period T and is piecewise continuos on $[0,T]$, then

$$\mathcal{L}(f)(s) = \frac{F_T(s)}{1-e^{-sT}} = \frac{\int_0^T e^{-st} f(t)\, dt}{1-e^{-sT}}$$

[1] If you want to really understand the Γ function then take Math 331 and then find time for a conversation with Dr. Smith.

Proof: We will use the unit step function to express the function in a format which permits us to use Theorem 6.4.4. Assume $dom(f) = [0, \infty)$

$$f(t) = f_T(t) + f_T(t-T)u(t-T) + f_T(t-2T)u(t-2T) + ...$$

This is sneaky in that $f_T(t-T) \neq 0$ only for $T < t < 2T$ and $f_T(t-2T) \neq 0$ only for $2t < t < 3t$. so the unit step functions just multiply by 1 amd are superfluous as these shifted f_T functions are already set-up to be zero most places. We want the unit step functions so we can use Theorem 6.4.4:

$$\begin{aligned}
\mathcal{L}(f)(s) &= \mathcal{L}(f_T)(s) + \mathcal{L}(f_T(t-T)u(t-T))(s) + ... \\
&= F_T(s) + e^{-sT}F_T(s) + e^{-2sT}F_T(s) + ... \\
&= \frac{F_t(s)}{1 - e^{-sT}}
\end{aligned}$$

for $|e^{-st}| < 1$. The theorem follows. □

Example 6.5.5. Problem: *$f_T(t) = e^t$ and periodic f has $T = 1$. Calculate the Laplace transform of f.*

Solution: *We have:*

$$\begin{aligned}
\mathcal{L}(f)(s) &= \frac{\int_0^1 e^{-st}e^t\,dt}{1 - e^{-s}} \\
&= \frac{1}{1 - e^{-s}} \int_0^1 e^{t(1-s)}\,dt \\
&= \frac{1}{1 - e^{-s}} \cdot \frac{1}{1 - s} e^{t(1-s)} \Big|_0^1 \\
&= \frac{1}{1 - e^{-s}} \cdot \frac{1}{1 - s}(e^{1-s} - 1) \\
&= \frac{1}{s - 1} \cdot \frac{e^s - e}{e^s - 1}.
\end{aligned}$$

Example 6.5.6. Problem: *let $f(t) = \begin{cases} \frac{\sin t}{t}, & if\ t \neq 0 \\ 1, & ift = 0 \end{cases}$. Find the Laplace transform of f.*

Solution: *Recall that*

$$\sin t = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n+1}}{(2n+1)!} \quad \Rightarrow \quad \frac{\sin t}{t} = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n}}{(2n+1)!}$$

It is not hard to see that f is of exponential order. Thus, we expect its Laplace transform exists. And in fact it can be shown that the Laplace transform of a series is the series of the Laplace transform of the terms. That is we can extend linearity of $\mathcal{L}$ to infinite sums provided the series is *well-behaved (need exponential order).*

$$\begin{aligned}
\mathcal{L}(f)(s) &= \mathcal{L}\left(\sum_{n=0}^{\infty}(-1)^n \frac{t^{2n}}{(2n+1)!}\right)(s) \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{1}{(2n+1)!}\mathcal{L}(t^{2n})(s) \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{1}{(2n+1)!}\frac{(2n)!}{s^{2n+1}} \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{1}{2n+1}\frac{1}{s^{2n+1}} \\
&= \tan^{-1}\left(\frac{1}{s}\right).
\end{aligned}$$

Definition 6.5.7.

The gamma function Γ is defined on $(0,\infty)$ as

$$\Gamma(t) := \int_0^{\infty} e^{-u}u^{t-1}\,du$$

Integration by parts shows that Γ has the property $\Gamma(t+1) = t\Gamma(t)$. Furthermore, notice that

$$\Gamma(1) = \int_0^{\infty} e^{-u}\,du = 1.$$

If $n \in \mathbb{Z}$ then $\Gamma(n+1) = n\Gamma(n) = ... = n!$. This means the gamma function is a continuous extension of the factorial ! Previously, we have utilized $\mathcal{L}(t^n)(s) = \frac{n!}{s^{n+1}}$. However, for any non-negative n, we have:

$$\mathcal{L}(t^n)(s) = \frac{\Gamma(n+1)}{s^{n+1}}.$$

Let's prove it, take $s > 0$ as usual,

$$\mathcal{L}(t^n)(s) = \int_0^{\infty} e^{-st}t^n\,dt = \int_0^{\infty} e^{-u}\frac{u^n}{s^n}\frac{1}{s}\,du = \frac{1}{s^{n+1}}\int_0^{\infty} e^{-u}u^n\,du = \frac{\Gamma(n+1)}{s^{n+1}}.$$

The discussion above serves to prove the following:

Theorem 6.5.8.

$$\mathcal{L}(t^n)(s) = \frac{\Gamma(n+1)}{s^{n+1}}$$

Remark: The gamma funcion is important to probability theory.

Example 6.5.9. $\mathcal{L}(t^{3.6})(s) = \frac{\Gamma(4.6)}{s^{4.6}}$.

Values of the Γ-function typically must be determined by some numerical method or table.

6.6 The Dirac delta device

Definition 6.6.1.

The Dirac delta function $\delta(t)$ is characterized by

1. $\delta(t) = \begin{cases} 0, & \text{if } t \neq 0 \\ \infty, & \text{otherwise} \end{cases}$

2. $\int_{-\infty}^{\infty} f(t)\delta(t)\,dt = f(0)$

for any function f that is continuous on some neighborhood of zero.

Technically $\delta(t)$ is a "generalized function" or better yet a "distribution". It was introduced by P.A.M. Dirac for physics, but was only later justified by mathematicians. This trend is also seen in recent physics, the physics community tends to do calculations that are not well-defined. Fortunately, physical intuition has guided them to not make very bad mistakes for the most part. Dirac later introduced something that came to be named "Dirac string" to described a quirk in the mathematics of the magnetic monopole. It took more than 20 years for the mathematics to really catch up and better explain the Dirac String in terms of a beautiful mathematical construction of fiber bundles. I digress! Anyway, we sometimes say "we have it down to a science", it would be better to say "we have it down to a math".

Remark 6.6.2.

Dirac delta functions in 3−dimensions work about the same: $\delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z)$, for $\mathbf{r} = (x, y, z)$. One application is to model the ideal point charge q at location $\mathbf{a}$ it has charge density

$$\rho(\mathbf{r}) = q\delta(\mathbf{r} - \mathbf{a})$$

see Griffith's Electrodynamics or take a junior-level electricity and magnetism course.

Heuristic Justification of $\delta(t)$: Impulse is defined to be the time integral of force experienced by some object over some short time interval $t_0 \to t_1$,

$$\text{Impulse} := \int_{t_0}^{t_1} F(t)\,dt = \int_{t_0}^{t_1} \frac{dP}{dt}\,dt = P(t_1) - P(t_0)$$

Since $F = \frac{dP}{dt}$, it is also equal to the change in momentum. You might think of a hammer striking a nail or a ball bouncing off a wall. A large force is exerted over a short time period.

You can imagine applying a greater force over a shorter time till in the limit you approach the notion of the delta function. The Dirac delta function can be used to model an impulse where we do not perhaps know the details of the impact, but we know it happened quickly and with a certain overall change in momentum. In such case, the $\delta(t)$ provides a useful idealization. Similarly when it is used to describe the charge densities, it provdies us a convenient mathematics for describing a localized source. I think a point charge is not really a point, but rather an extended (although tiny) body. We don't know the details of such tiny things, or if we do they are complicated. In such cases the Dirac delta function provides a useful mathematical idealization.

Proposition 6.6.3.

$$\mathcal{L}(\delta(t-a))(s) = e^{-as}$$

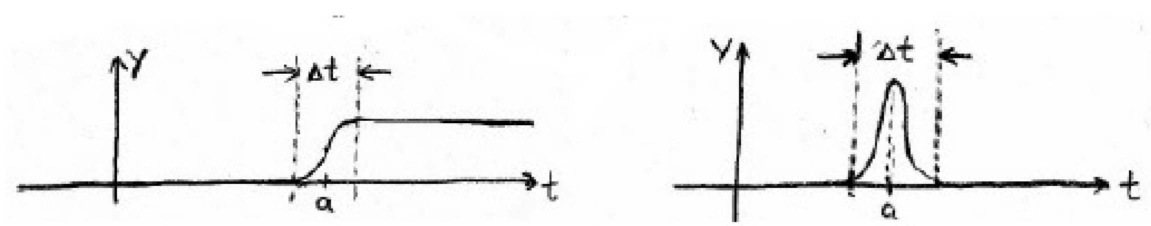
Proof: Integrals with δ-functions are easy; integration becomes evaluation:

$$\mathcal{L}(\delta(t-a))(s) = \int_0^\infty e^{-st}\delta(t-a)\,dt = e^{-st}|_{t-a=0} = e^{-as}.$$

Consider the graph of the unit step function $u(t-a)$. The unit step is constant everywhere except at $t = a$, where it jumps 1 in zero-time. This suggests,

$$\frac{d}{dt}(u(t-a)) = \delta(t-a).$$

Of course, there is something exotic here, $u(t-a)$ is not even continuous, what right have we to differentiate it? Of course the result $\delta(t-a)$ is not a function and the $\frac{d}{dt}$ here is a more general idea that the one seen in Calculus I. $\square$



Remark 6.6.4.

In reality applications typically have functions more like the picture above. As $\delta t \to 0$, we obtain the unit step function and $\delta(t)$. We use $u(t-a)$ amd $\delta(t-a)$ necaise we either do not know or do not care the details of what happens close to time $t = a$.

Example 6.6.5. Problem: *consider a mass spring system with no friction and $m = 1$ and spring constant $k = 1$ initially at rest $x(0) = 0$ and $x'(0) = 0$ hit by hammer at $t = 0$. Model the force by $F(t) = \delta(t)$. Find the resulting motion.*

Solution: *We apply Newton's Second Law. This can be described by (no friction and $m = k = 1$)*

$$x'' + x = \delta(t).$$

Suppose the system has . Then

$$s^2X + X = \mathcal{L}(\delta(t))(s) = e^0 = 1 \quad \Rightarrow \quad X = \frac{1}{s^2+1} \quad \Rightarrow \quad x(t) = \sin t.$$

Notice that while $x(0) = 0$, $x'(0) = 1 \neq 0$. This is to be expected as $x'(0^-) = 0$ while $x'(0^+)$ Since $\Delta P = m\Delta V = 1$. So the velocity has to change very quickly.

Example 6.6.6. Problem: *solve:* $x'' + 9x = 3\delta(t-\pi)$, *where* $x(0) = 1$ *and* $x'(0) = 0$.

Solution: *We have:* $s^2X - s + 9X = 3e^{-\pi s}$, *which implies*

$$(s^2+9)X = 3e^{-\pi s} + s \quad \Rightarrow \quad X = 3e^{-\pi s}\frac{1}{s^2+9} + \frac{s}{s^2+9}$$

Thus we have

$$\begin{aligned} x(t) &= \mathcal{L}^{-1}(X)(t) \\ &= \mathcal{L}^{-1}\left(\frac{3e^{-\pi s}}{s^2+9}\right)(t) + \mathcal{L}^{-1}\left(\frac{s}{s^2+9}\right)(t) \\ &= \sin(3(t-\pi))u(t-\pi) + \cos 3t \\ &= \begin{cases} \cos 3t, & \text{if } \ t < \pi \\ \cos 3t - \sin 3t, & \text{if } t > \pi \end{cases}. \end{aligned}$$

Example 6.6.7. Problem: *Solve* $y'' - 2y' + y = u(t-1)$, *where* $y(0) = 0$ *and* $y'(0) = 1$.

Solution: *We have* $s^2Y - 1 - 2sY + Y = \frac{e^{-s}}{s}$, *which implies*

$$(s^2 - 2s + 1)Y = 1 + \frac{e^{-s}}{s}$$

Thus we have:

$$Y = \frac{1}{s^2-2s+1} + \frac{e^{-s}}{s(s^2-2s+1)}$$

Define $F(s)$ *as follows and perform the partial-fractal-decomposition:*

$$F(s) = \frac{1}{s(s^2-2s+1)} = \frac{1}{s} - \frac{1}{s-1} + \frac{1}{(s-1)^2}.$$

Therefore, $f(t) = 1 - e^t + te^t$. *Hence* $\mathcal{L}^{-1}(F(s)e^{-s})(t) = (1 - e^{t-1} + (t-1)e^{t-1})u(t-1)$. *Thus, noting* $s^2 - 2s + 1 = (s-1)^2$ *to simplify the formula for* Y,

$$\begin{aligned} \mathcal{L}^{-1}(Y)(t) &= \mathcal{L}^{-1}\left(\frac{1}{(s-1)^2}\right) + \mathcal{L}^{-1}\left(e^{-s}F(s)\right) \\ &= te^t + (1 - e^{t-1} + (t-1)e^{t-1})u(t-1). \end{aligned}$$

Therefore, the solution is given by:

$$y(t) = te^t + (1 - e^{t-1}(t-2))u(t-1).$$

Example 6.6.8. Problem: *solve* $y'' + 5y' + 6y = u(t-1)$, *with* $y(0) = 2$, $y'(0) = 1$.

Solution: *We have* $s^2Y - 2s - 1 + 5(sY - 2) + 6Y = \frac{e^{-s}}{s}$. *Thus*

$$(s^2+5s+6)Y = 2s + 1 + 10 + \frac{e^{-s}}{s} \quad \Rightarrow \quad Y = \frac{2s+11}{s^2+5s+6} + \frac{e^{-s}}{s(s^2+5s+6)}$$

Note that

$$\frac{2s+11}{s^2+5s+6} = \frac{7}{s+2} - \frac{5}{s+3}$$

Also, define and decompose $F(s)$ as follows:

$$F(s) = \frac{1}{s(s^2+5s+6)} = \frac{1/6}{s} + \frac{-1/2}{s+2} + \frac{1/3}{s+3}$$

Thus $f(t) = \frac{1}{6} - \frac{1}{2}e^{-2t} + \frac{1}{3}e^{-3t}$. Therefore,

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}\left(\frac{7}{s+2} - \frac{5}{s+3}\right)(t) + \mathcal{L}^{-1}(F(s)e^{-s})(t) \\ &= 7e^{-2t} - 5e^{-3t} + f(t-1)u(t-1) \\ &= 7e^{-2t} - 5e^{-3t} + \left(\frac{1}{6} - \frac{1}{2}e^{-2(t-1)} + \frac{1}{3}e^{-3(t-1)}\right)u(t-1) \end{aligned}$$

Example 6.6.9. Problem: *solve $y'' - 7y' + 12y = 12u(t-4)$, where $y(0) = 1$ and $y'(0) = 4$.*

Solution: *We have $s^2Y - s - 4 - 7(sY - 1) + 12Y = \frac{12}{s}e^{-4s}$. Therefore,*

$$(s^2 - 7s + 12)Y = s - 3 + 12\frac{e^{-4s}}{s}$$

Solve for Y, and implicitly define $F(s)$ by the rightmost equality

$$Y = \frac{s-3}{s^2-7s+12} + \frac{12e^{-4s}}{s(s^2-7s+12)} = \frac{1}{s-4} + F(s)e^{-4s}.$$

Partial fractions for $F(s)$ is found after a short calculation:

$$F(s) = \frac{12}{s(s-3)(s-4)} = \frac{1}{s} - \frac{4}{s-3} + \frac{3}{s-4}.$$

Thus,

$$f(t) = \mathcal{L}^{-1}(F)(t) = 1 - 4e^{3t} + 3e^{4t}.$$

Now we assemble the solution:

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}(Y)(t) \\ &= \mathcal{L}^{-1}\left(\frac{1}{s-4}\right)(t) + \mathcal{L}^{-1}(F(s)e^{-4s})(t) \\ &= e^{4t} + f(t-4)u(t-4) \\ &= e^{4t} + \left(1 - 4e^{3(t-4)} + 3e^{4(t-4)}\right)u(t-4). \end{aligned}$$

Example 6.6.10. Problem: *solve $y'' - 7y' + 12y = 12u(t-4)$, where $y(0) = 1$, and $y'(0) = \frac{1}{2}$*

Solution: *We have $s^2Y - s - \frac{1}{2} - 7(sY - 1) + 12Y = \frac{12e^{-4s}}{s}$, which implies*

$$(s^2 - 7s + 12)Y = s + \frac{1}{2} - 7 + \frac{12e^{-4s}}{s} = s - \frac{13}{2} + \frac{12e^{-4s}}{s}$$

Solve for Y,

$$Y = \frac{s - 13/2}{s^2-7s+12} + \frac{12e^{-4s}}{s(s^2-7s+12)}.$$

Note that

$$\frac{s-13/2}{(s-3)(s-4)} = \frac{7/2}{s-3} - \frac{5/2}{s-4}.$$

Also, define and decompose $F(s)$ into basic rational functions:

$$F(s) = \frac{12}{s(s-3)(s-4)} = \frac{1}{s} - \frac{4}{s-3} + \frac{3}{s-4}.$$

Thus $f(t) = 1 - 4e^{3t} + 3e^{4t}$. Therefore,

$$y(t) = \frac{7e^{3t}}{2} - \frac{5e^{4t}}{2} + \left(1 - 4e^{3(t-4)} + 3e^{4(t-4)}\right) u(t-4).$$

6.7 Convolution

It turns out that if we want to map the product of functions in the s-domain[2] to a corresponding product in the t-domain then the following convoluted product is what we need:

Definition 6.7.1.

Let $f(t)$ and $g(t)$ be piecewise continuous on $[0,\infty)$. The convolution of f and g is denoted $f * g$ and is defined by

$$(f*g)(t) = \int_0^t f(t-v)g(v)\,dv$$

Example 6.7.2. *The convolution of t and t^2 on $[0,\infty)$ is*

$$t * t^2 = \int_0^t (t-v)v^2\,dv = \int_0^t (tv^2 - v^3)\,dv = \left(\frac{1}{3}tv^3 - \frac{1}{4}v^4\right)\Big|_0^t = \frac{1}{12}t^4.$$

Theorem 6.7.3.

Given piecewise continuous functions f, g, h on $[0,\infty)$, we have

1. $f * g = g * f$
2. $f * (g + h) = f * g + f * h$
3. $f * (g * h) = (f * g) * h$
4. $f * 0 = 0$

Proof: left to reader. □

The theorem below is the whole reason for defining such a thing as a "convolution". You could derive the formula for the convolution by working backwards from this theorem.

[2]this is made precise by the Convolution Theorem 6.7.4

Theorem 6.7.4. *Covolution Theorem*

Given f, g piecewise continuous on $[0, \infty)$ and of exponential order α with Laplace transforms

$$\mathcal{L}(f)(s) = F(s) \quad \& \quad \mathcal{L}(g)(s) = G(s)$$

then,

$$\mathcal{L}(f * g)(s) = F(s)G(s)$$

Or in other words,

$$\mathcal{L}^{-1}(FG)(t) = (f * g)(t).$$

Proof: is essentially just the calculation below:

$$\begin{aligned}
\mathcal{L}(f * g)(s) &= \int_0^\infty e^{-st}(f * g)(t)\,dt \\
&= \int_0^\infty e^{-st}\left(\int_0^t f(t-v)g(v)\,dv\right)\,dt \\
&= \int_0^\infty e^{-st}\left(\int_0^\infty u(t-v)f(t-v)g(v)\,dt\right)\,dt \\
&= \int_0^\infty g(v)\left(\int_0^\infty e^{-st}u(t-v)f(t-v)\,dt\right)\,dv \\
&= \int_0^\infty g(v)\mathcal{L}(u(t-v)f(t-v))(s)\,dv \\
&= \int_0^\infty g(v)e^{-sv}F(s)\,dv \\
&= F(s)\int_0^\infty e^{-sv}g(v)\,dv \\
&= F(s)G(s). \qquad \square
\end{aligned}$$

Example 6.7.5. Problem: *solve $y'' + y = g(t)$ with $y(0) = 0$ and $y'(0) = 0$.*

Solution: *the Laplace transform yields $s^2Y + Y = G$. This implies*

$$Y(s) = \frac{1}{s^2+1}G(s)$$

Using the convolution theorem, we obtain,

$$y(t) = \mathcal{L}^{-1}(Y)(t) = \mathcal{L}^{-1}\left(\frac{1}{s^2+1}G(s)\right)(t) = \left(\mathcal{L}^{-1}\left(\frac{1}{s^2+1}\right) * \mathcal{L}^{-1}(G)\right)(t) = \sin t * g(t)$$

This is an integral solution to the differential equation. This result is quite impressive, notice it works for any piecewise continuos forcing term of exponential order.

Example 6.7.6. **Problem:** *solve: $y'' + y = \cos t$, with $y(0) = y'(0) = 0$.*

Solution: *We make use of the previous example. We found that:*

$$\begin{aligned} y &= \int_0^t \sin(t-v)\cos v\, dv \\ &= \int_0^t (\sin t \cos^2 v - \sin v \cos t \cos v)\, dv \\ &= \sin t \int_0^t \frac{1}{2}(1+\cos 2v)\, dv - \cos t \int_0^t \sin v \cos v\, dv \\ &= \frac{1}{2}\sin t(t + 1/2 \sin 2t) - \cos t\left(\frac{1}{2}\sin^2 t\right) \\ &= \frac{1}{2} t \sin t. \end{aligned}$$

Remark 6.7.7.

This gives us yet another method to explain the presence of the factor "t" in y_p when there is overlap. Here $y_1 = \cos t$ and $y_2 = \sin t$ are the fundamental solution, clearly $\cos t = g(t)$ overlaps.

Example 6.7.8. **Problem:** *solve $y'' - y = g(t)$ with $y(0) = 1$ and $y'(0) = 1$. Assuming that $g(t)$ has well-defined Laplace transform G.*

Solution: *We have $s^2Y - sy(0) - y'(0) - Y = G$, which implies*

$$Y(s) = \frac{1}{s-1} + \frac{1}{s^2-1}G(s)$$

Notice by partial fractions $\frac{1}{s^2-1} = \frac{1/2}{s-1} - \frac{1/2}{s+1}$. Therefore,

$$\mathcal{L}^{-1}\left(\frac{1}{s^2-1}\right)(t) = \frac{1}{2}\mathcal{L}^{-1}\left(\frac{1}{s-1}\right)(t) - \frac{1}{2}\mathcal{L}^{-1}\left(\frac{1}{s+1}\right)(t) = \frac{1}{2}(e^t - e^{-t})$$

Hence using the convolution theorem,

$$y(t) = e^t + \sinh(t) * g(t) = e^t + \int_0^t \sinh(t-v)g(v)\, dv.$$

Example 6.7.9. Problem: *use convolution theorem to find* $\mathcal{L}^{-1}\left(\frac{1}{(s^2+1)^2}\right)$.

Solution: *We have* $\frac{1}{(s^2+1)^2} = \frac{1}{s^2+1}\frac{1}{s^2+1} = F(s)G(s)$*. Now* $\mathcal{L}^{-1}(F)(t) = \sin t = f(t) = g(t) = \mathcal{L}^{-1}(G)(t)$.

$$\begin{aligned}\mathcal{L}^{-1}\left(\frac{1}{(s^2+1)^2}\right)(t) = \mathcal{L}^{-1}(FG)(t) &= f(t) * g(t) \\ &= \int_0^t f(t-v)g(v)\,dv \\ &= \int_0^t \sin(t-v)\sin v\,dv \\ &= \frac{1}{2}\sin^3 t - \cos t\left(\frac{1}{2}t - \frac{1}{4}\sin 2t\right) \\ &= \frac{1}{2}\sin t - \frac{1}{2}t\cos t.\end{aligned}$$

Remark 6.7.10.

The convolution theorem allows us to unravel many inverse Laplace transforms in a slick way. In the other direction it is not as useful since as a starting point you need to identify some convolution in t. Unless your example is very special it is unlikely the convolution theorem will be useful in taking the Laplace transform. I should also mention, the concept of a **transfer function** and the associated linear systems analysis is an interesting topic which is intertwined with the convolution technique.

6.8 practice problems

PP 310 Calculate the Laplace transform of $f(t) = t$ from the definition of the Laplace transform. That is, calculate $\mathcal{L}\{t\}(s) = \int_0^\infty te^{-st}dt$.

PP 311 Calculate $\mathcal{L}\{te^{3t}\}(s)$.

PP 312 Let $f(t) = \sin t$ for $0 \leq t \leq \pi$ and $f(t) = 0$ for $t > \pi$. Calculate $\mathcal{L}\{f(t)\}(s)$.

PP 313 Calculate $\mathcal{L}\{e^{3t}\sin(6t) - t^3 + e^t\}(s)$

PP 314 Calculate $\mathcal{L}\{t^4 - t^2 - t + \sin(\sqrt{2}t)\}(s)$

PP 315 Calculate $\mathcal{L}\{2t^2e^{-t}\}(s)$

PP 316 Calculate $\mathcal{L}\{t^2e^{3t} + e^{-2t}\sin(2t)\}(s)$

PP 317 Calculate $\mathcal{L}\{\sin(3t)\cos(3t)\}(s)$

PP 318 Calculate $\mathcal{L}\{\cos^3(t)\}(s)$

PP 319 Derive $\mathcal{L}\{t^n\}(s) = \dfrac{n!}{s^{n+1}}$

PP 320 Derive $\mathcal{L}\{\cosh(bt)\}(s) = \dfrac{s}{s^2 - b^2}$

PP 321 Calculate $\mathcal{L}^{-1}\left\{\dfrac{6}{(s-1)^4}\right\}(t)$.

PP 322 Calculate $\mathcal{L}^{-1}\left\{\dfrac{2}{s^2+4}\right\}(t)$.

PP 323 Calculate $\mathcal{L}^{-1}\left\{\dfrac{s+1}{s^2+2s+10}\right\}(t)$.

PP 324 Calculate $\mathcal{L}^{-1}\left\{\dfrac{1}{s^5}\right\}(t)$.

PP 325 Let $F(s) = \dfrac{3s-15}{2s^2-4s+10}$. Calculate $f = \mathcal{L}^{-1}\{F\}$.

PP 326 Let $F(s) = \dfrac{6s^2-13s+2}{s(s-1)(s-6)}$. Calculate $f = \mathcal{L}^{-1}\{F\}$.

PP 327 Let $F(s) = \dfrac{s+11}{(s-1)(s+3)}$. Calculate $f = \mathcal{L}^{-1}\{F\}$.

PP 328 Let $F(s) = \dfrac{7s^3-2s^2-3s+6}{s^3(s-2)}$. Calculate $f = \mathcal{L}^{-1}\{F\}$.

PP 329 Given $s^2F(s) + sF(s) - 6F(s) = \dfrac{s^2+4}{s^2+5}$ calculate $f = \mathcal{L}^{-1}\{F\}$.

PP 330 Let $F(s) = \ln\left(\dfrac{s+2}{s-5}\right)$. Calculate $f = \mathcal{L}^{-1}\{F\}$.

PP 331 Let $F(s) = \tan^{-1}\left(\dfrac{1}{s}\right)$. Calculate $f = \mathcal{L}^{-1}\{F\}$.

PP 332 Solve $y'' - 2y' + 5y = 0$ with $y(0) = 2$ and $y'(0) = 4$ via the Laplace transform technique.

PP 333 Solve $y''+6y'+5y = 12e^t$ with $y(0) = -1$ and $y'(0) = 7$ via the Laplace transform technique.

PP 334 Solve $w'' + w = t^2 + 2$ with $w(0) = 1$ and $w'(0) = -1$ via the Laplace transform technique.

PP 335 Solve $y'' - 4y = 4t - 8e^{-2t}$ with $y(0) = 0$ and $y'(0) = 5$ via the Laplace transform technique.

PP 336 Solve $y'' + 3ty' - 6y = 1$ with $y(0) = 0$ and $y'(0) = 0$ via the Laplace transform technique.

PP 337 Solve $y'' + y = t$ with $y(\pi) = 0$ and $y'(\pi) = 0$ via the Laplace transform technique.

PP 338 Let $g(t) = \begin{cases} 0, & 0 < t < 2 \\ t+1, & 2 < t \end{cases}$. Calculate $G(s)$.

PP 339 Let $g(t) = \begin{cases} 0, & 0 < t < 1 \\ 2, & 1 < t < 2 \\ 1, & 2 < t < 3 \\ 3, & 3 < t \end{cases}$. Calculate $G(s)$.

PP 340 Let $g(t) = \begin{cases} 0, & t < 1 \\ t-1, & 1 < t < 2 \\ 3-t, & 2 < t < 3 \\ 0, & 3 < t \end{cases}$. Calculate $G(s)$.

PP 341 Let $G(s) = \dfrac{e^{-3s}}{s^2}$. Calculate $g(t)$.

PP 342 Calculate $\mathcal{L}^{-1}\left\{\dfrac{e^{-2s} - 3e^{-4s}}{s+2}\right\}(t)$.

PP 343 Calculate $\mathcal{L}^{-1}\left\{\dfrac{e^{-s}}{s^2+4}\right\}(t)$.

PP 344 Solve $y'' + 4y' + 4y = u(t-\pi) - u(t-2\pi)$ with $y(0) = 0$ and $y'(0) = 0$ via the Laplace transform technique.

PP 345 Solve $y'' + 5y' + 6y = g(t)$ given $y(0) = 0$ and $y'(0) = 2$ where $g(t) = \begin{cases} 0, & 0 < t < 1 \\ t, & 1 < t < 5 \\ 1, & 5 < t \end{cases}$.

PP 346 Solve $y'' - y = u(t-1) - u(t-4)$ given $y(0) = 0$ and $y'(0) = 2$.

PP 347 Calculate $\int_{-\infty}^{\infty} (t^2 - 1)\delta(t)dt$.

PP 348 Calculate $\int_{-\infty}^{\infty} e^{3t}\delta(t)dt$.

PP 349 Calculate $\int_{-\infty}^{\infty} \sin(3t)\delta\left(t - \frac{\pi}{2}\right) dt$.

PP 350 Calculate $\int_{-\infty}^{\infty} e^{-2t}\delta(t+1)dt$.

PP 351 Calculate $\mathcal{L}\{\delta(t-1) - \delta(t-3)\}(s)$.

PP 352 Calculate $\mathcal{L}\{\delta(t-\pi)\sin t\}(s)$.

PP 353 Solve $w'' + w = \delta(t-\pi)$ where $w(0) = 0$ and $w'(0) = 0$.

PP 354 Solve $y'' + y = 4\delta(t-2) + t^2$ given $y(0) = 0$ and $y'(0) = 2$.

PP 355 A hammer hits a spring mass system at time $t = \pi/2$ and thus Newton's Second Law gives

$$\frac{d^2x}{dt^2} + 9x = -3\delta(t - \pi/2)$$

with $x(0) = 1$ and $x'(0) = 0$ since the spring is initially stretched to 1-unit and released from rest. Calculate the equation of motion and explain what happens after the hammer hits the spring at time $t = \pi/2$.

PP 356 Calculate the Laplace transforms of the following functions

(a.) $f(t) = \sin(t)\cos(2t) + \sin^2(3t)$

(b.) $f(t) = e^t u(t-3) + \sin(t)u(t-6)$

PP 357 Calculate the Laplace transforms of the following functions

(a.) $f(t) = \begin{cases} t, & 0 \leq t \leq 2 \\ \sin(t) & t > 2 \end{cases}$.

(b.) $f(t) = te^{-2t} + t\sin(t)$

PP 358 Compute the inverse Laplace transforms of $F(s) = \dfrac{3s+9}{s^2 - 8s + 7}$

PP 359 Compute the inverse Laplace transform of $F(s) = \dfrac{e^{-2s}}{s(s^2 + 6s + 13)}$

PP 360 Compute the inverse Laplace transform of $F(s) = \dfrac{4s}{s^4 - 1}$

PP 361 Solve the following differential equations with the given initial conditions by the method of Laplace transforms.

(a.) $y'' + y' - 2y = 0$ where $y(0) = 2$ and $y'(0) = 1$

(b.) $y'' - 2y' + y = \delta(t-2)$ where $y(0) = 1$ and $y'(0) = 0$

PP 362 Solve $y'' - 8y' + 7y = u(t-2)$ where $y(0) = 0$ and $y'(0) = 0$ by the method of Laplace transforms.

PP 363 Solve $y'' - 8y' + 7y = u(t-2) + u(t-4)$ where $y(0) = 0$ and $y'(0) = 0$ by the method of Laplace transforms.

PP 364 Solution of IVP with periodic forcing functions.

(a.) find the Laplace transform of the periodic function f where $T = 2a$ and we define $f(t) = 1$ for $0 < t < a$ and $f(t) = 0$ for $a \leq t \leq 2a$. This is the square wave pictured in Problem 25 of Nagel Saff and Snider, page 422 of §7.6. (5th edition, you might need to look around given the current edition)

(b.) solve $y'' + 3y' + 2y = f(t)$ for $t > 0$ given $y(0) = y'(0) = 0$.

PP 365 A spring with stiffness $k = 4$ is attached to a mass $m = 1$ and oscillates in one-dimensional motion such that it has $x(0) = 1$ and $x'(0) = 1$. Is it possible to strike the mass / spring system with a hammer such that the system is motionless after the strike ? Assume an idealized hammer which produces a force $F(t) = J_o\delta(t-a)$, you are free to adjust J_o and a as needed.

PP 366 (Ritger & Rose section 9-6 problem 1a) Use convolution to find the inverse Laplace transform of $\dfrac{1}{s^2(s-a)}$ for $a \neq 0$.

PP 367 Find an integral solution of $y'' + y = g$ via Laplace transforms and convolution. You may assume g is an integrable function of time t.

PP 368 (Ritger & Rose pg. 302 section 9-8) Suppose $L[y] = f$ is a second order linear system. If the possible inputs (we use a complex notation to treat sines and cosines at once) are given by $f(t) = ce^{i\omega t}$ for $c \in \mathbb{C}$ and $\omega \in \mathbb{R}$ then **show** that the output is given by

$$y(t) = H(i\omega)ce^{i\omega t} + y_t(t)$$

where $y_t(t) \to 0$ as $t \to \infty$ (y_t is the *transient solution*). The function $H(i\omega)$ is called the **frequency-response function** for the system. Notice that we can express

$$H(i\omega) = A(\omega)e^{i\phi(\omega)}$$

The factor $A(\omega)$ is the **amplification factor** for the system whereas $\phi(\omega)$ is the **phase lag**. **Find formulas** for $A(\omega) \in (0, \infty)$ and $\phi(\omega)$ in the particular cases:

(a.) $H(s) = \dfrac{1}{s^2 + 5s + 6}$

(b.) $H(s) = \dfrac{1}{s^2 + s + 1}$

(c.) $H(s) = \dfrac{1}{s^2 + s}$

PP 369 Consider $L[y] = (D - 2)(D^2 + 4D + 5)[y] = f$ where $D = d/dt$. Find:

(a.) green's function $K(u, t)$ (see my notes for the meaning of this),

(b.) transfer function $H(s)$ and $h(t)$,

(c.) an integral solution of $L[y] = f$ subject $y(0) = y'(0) = y''(0) = 0$ for $f(t) = t^2 \cos(t)$. **NOTE: DO NOT DO THIS INTEGRAL, THIS IS WHAT IS MEANT BY "INTEGRAL" SOLUTION, IT IS THE ANSWER REDUCED TO AN INTEGRAL**

PP 370 Calculate the Laplace transforms of the following functions using the table of basic Laplace transforms plus possibly the given Theorems and trigonometry.

(a.) $f(t) = \begin{cases} t, & 0 \leq t \leq 2 \\ \sin(t) & t > 2 \end{cases}$.

(b.) $f(t) = te^{-2t} + t\sin(t)$

PP 371 Compute the inverse Laplace transforms of,

(a.) $F(s) = \dfrac{3s + 9}{s^2 - 8s + 7}$

(b.) $F(s) = \dfrac{e^{-2s}}{s(s^2 + 6s + 13)}$

(c.) $F(s) = \dfrac{4s}{s^4 - 1}$

PP 372 Solve the following differential equations with the given initial conditions by the method of Laplace transforms.

(a.) $y'' + y' - 2y = 0$ where $y(0) = 2$ and $y'(0) = 1$

(b.) $y'' - 2y' + y = \delta(t - 2)$ where $y(0) = 1$ and $y'(0) = 0$

PP 373 Solve the following differential equations with the given initial conditions by the method of Laplace transforms.

(a.) $y'' - 8y' + 7y = u(t - 2)$ where $y(0) = 0$ and $y'(0) = 0$

(b.) $y'' - 8y' + 7y = u(t-2) + u(t-4)$ where $y(0) = 0$ and $y'(0) = 0$

PP 374 Solution of IVP with periodic forcing functions.

(a.) find the Laplace transform of the periodic function f where $T = 2a$ and we define $f(t) = 1$ for $0 < t < a$ and $f(t) = 0$ for $a \leq t \leq 2a$. This is the square wave pictured in Problem 25 of Nagel Saff and Snider, page 422 of §7.6.

(b.) solve $y'' + 3y' + 2y = f(t)$ for $t > 0$ given $y(0) = y'(0) = 0$.

PP 375 Let $f(t) = \begin{cases} \sin(t) & 0 \leq t \leq 2 \\ e^t & t > 2 \end{cases}$. Calculate the Laplace transform of f.

PP 376 Suppose $F(s) = \dfrac{72s}{s^4 - 81}$. Calculate the inverse Laplace transform of $F(s)$.

PP 377 Solve $y'' + 6y' + 13y = u(t-1)$ given $y(0) = 1$ and $y'(0) = 3$

Chapter 7

the series solution technique

Series techniques have been with us a long time now. Founders of calculus worked with series in a somewhat careless fashion and we will do the same here. The wisdom of nineteenth century analysis is more or less ignored in this work. In short, I am not too worried about the interval of convergence in these notes. This is of course a dangerous game, but the density of math majors permits no other. I'll just make this comment: the series we find generally represent a function of interest only locally. Singularities prevent us from continuing the expansion past some particular point.

It doesn't concern this course too much, but perhaps it's worth mentioning: much of the work we see here arose from studying complex differential equations. The results for ordinary points were probably known by Euler and Lagrange even took analyticity as a starting point for what he thought of as a "function". The word "analytic" should be understood to mean that there exists a power series expansion representing the function near the point in question. There are functions which are not analytic and yet are smooth ($f(f(x)) = \sin(x)$ defines such a function, see the math stack for more). Logically, functions need not be analytic. However, most nice formulas do impart analyticity at least locally.

Fuchs studied complex differential equations as did Weierstrauss, Cauchy, Riemann and most of the research math community of the nineteenth century. Fuchasian theory of DEqns dealt with the problem of singularities and there was (is) a theory of majorants due to Weierstrauss which was concerned with how singularities appear in solutions. In particular, the study of moveable singularities, the process of what we call *analytic continuation* was largely solved by Fuchs. However, his approach was more complicated than the methods we study. Frobenous proposed a method which clarified Fuch's work and we use it to this day. Read Hille's masterful text about differential equations in the complex plane for a more complete history. My point to you here is simply this: what we do here did not arise from the study of the real-valued problems we study alone. To really understand the genesis of this material you must study complex differential equations. We don't do this since complex variables are not a prerequisite for this course.

The calculations in this chapter can be challenging. However, the power series approximation is one of our most flexible tools for mathematical modelling and it is most certainly worth understanding. If you compare these notes with Ritger & Rose then you'll notice that I have not covered too deeply the sections towards the end of Chapter 7; Bessel, Legendre, and the hypergeometric equations are interesting problems, but it would take several class periods to absorb the material and I think it better to spend our time on breadth. My philosophy is that once you've taken this course you ought to be ready to do further study on those sections.

7.1 calculus of series

I begin with a brief overview of terminology and general concepts about sequences and series. We will not need all of this, but I think it is best to at least review the terms as to recover as much as we can from your previous course work.
A sequence in S is a function $a : \{n_o, n_o + 1, n_o + 2, \dots\} \to S$ where we usually denote $a(n) = a_n$ for all $n \in \mathbb{Z}$ with $n \geq n_o$. Typically $n_o = 0$ or $n_o = 1$, but certainly it is interesting to consider other initial points for the domain. If $a_n \in \mathbb{R}$ for all n then we say $\{a_n\}$ is a sequence of real numbers. If $a_n \in \mathbb{C}$ for all n then we say $\{a_n\}$ is a complex sequence. If $\mathcal{F}$ is a set of functions and $a_n \in \mathcal{F}$ for all n then we say $\{a_n\}$ is sequence of functions. If the codomain for a sequence has an operation such as addition or multiplication then we can add or multiply such sequences by the usual pointwise defined rules; $(ab)_n = a_n b_n$ and $(a + b)_n = a_n + b_n$. In addition, we can define a **series** $s = a_{n_o} + a_{n_1} + \cdots$ in S as follows:

$$s = \lim_{n\to\infty} \sum_{k=n_o}^{n} a_k$$

provided the limit above exists. In other words, the series above exists iff the sequence of partial sums $\{a_{n_o}, a_{n_o} + a_{n_1}, a_{n_o} + a_{n_1} + a_{n_2}, \dots\}$ converges. When the sequence of partial sums converges then the series is likewise said to converge and we can denote this by $s = \sum_{k=n_o}^{\infty} a_k$. You should remember studying the convergence of such series for a few weeks in your second calculus course. Perhaps you will be happy to hear that convergence is not the focus of our study in this chapter.

A **power function** is a function with formula $f(x) = x^n$ for some $n \in \mathbb{R}$. A **power series** is a series formed from adding together power functions. However, traditionally the term **power series** is reserved for series constructed with powers from $\mathbb{N} \cup \{0\}$. Equivalently we can say a **power series** is a function which is defined at each point by a series;

$$f(x) = \sum_{k=0}^{\infty} c_k (x - a)^k = c_o + c_1(x - a) + c_2(x - a)^2 + \cdots$$

The constants $c_o, c_1, c_2, \dots$ are fixed and essentially define f uniquely once the center point a is given. The domain of f is understood to be the set of all real x such that the series converges. Given that $f(x)$ is a power series it is a simple matter to compute that

$$c_o = f(a), \quad c_1 = f'(a), \quad c_2 = \frac{1}{2} f''(a), \quad \dots, c_k = \frac{1}{k!} f^{(k)}(a).$$

Incidentially, the result above shows that if $\sum_{k=0}^{\infty} b_k (x - a)^k = \sum_{k=0}^{\infty} c_k (x - a)^k$ then $b_k = c_k$ for all $k \geq 0$ since both power series define the same derivatives and we know derivatives are single-valued when they exist. This result is called **equating coefficients** of power series, we will use it many times.

The domain of a power series is somewhat boring. Recall that there are three possibilities:

1. $dom(f) = \{a\}$

2. $dom(f) = \{x \in \mathbb{R} \mid |x - a| \leq R\}$ for some radius $R > 0$.
3. $dom(f) = (-\infty, \infty)$

The constant R is called the **radius of convergence** and traditionally we extend it to all three cases above with the convention that for case (1.) $R = 0$ whereas for case (3.) $R = \infty$.

Given a function on $\mathbb{R}$ we can sometimes replace the given formula of the function with a power series. If it is possible to write the formula for the function f as a power series centered at x_o in some open set around x_o then we say f is **analytic at** x_o. When it is possible to write $f(x)$ as a single power series for all $x \in \mathbb{R}$ then we say f is **entire**. A function is called **smooth** at x_o if derivatives of arbitrary order exist for f at x_o. Whenever a function is smooth at x_o we can calculate $T(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_o)}{n!}(x - x_o)^n$ which is called the **Taylor series** of f centered at x_o. However, there are functions for which the series $T(x) \neq f(x)$ near x_o. Such a function is said to be **non-analytic**. If $f(x) = T(x)$ for all x close to x_o then we say f is analytic at x_o. This question is not treated in too much depth in most calculus II courses. It is much harder to prove a function is analytic than it is to simply compute a Taylor series. We again set-aside the issue of analyticity for a later course where analysis is the focus. We now turn our focus to the computational aspects of series.

If $f : U \subseteq \mathbb{R} \to \mathbb{R}$ is analytic at $x_o \in U$ then we can write

$$f(x) = f(x_o) + f'(x_o)(x - x_o) + \frac{1}{2}f''(x_o)(x - x_o)^2 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_o)}{n!}(x - x_o)^n$$

We could write this in terms of the operator $D = \frac{d}{dt}$ and the evaluation of $t = x_o$

$$f(x) = \left[\sum_{n=0}^{\infty} \frac{1}{n!}(x - t)^n D^n f(t)\right]_{t=x_o} =$$

I remind the reader that a function is called **entire** if it is analytic on all of $\mathbb{R}$, for example e^x, $\cos(x)$ and $\sin(x)$ are all entire. In particular, you should know that:

$$e^x = 1 + x + \frac{1}{2}x^2 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}x^n$$

$$\cos(x) = 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4 \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}x^{2n}$$

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1}$$

Since $e^x = \cosh(x) + \sinh(x)$ it also follows that

$$\cosh(x) = 1 + \frac{1}{2}x^2 + \frac{1}{4!}x^4 \cdots = \sum_{n=0}^{\infty} \frac{1}{(2n)!}x^{2n}$$

$$\sinh(x) = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 \cdots = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!}x^{2n+1}$$

The geometric series is often useful, for $a, r \in \mathbb{R}$ with $|r| < 1$ it is known

$$a + ar + ar^2 + \cdots = \sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$$

This generates a whole host of examples, for instance:

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots$$

$$\frac{1}{1-x^3} = 1 + x^3 + x^6 + x^9 + \cdots$$

$$\frac{x^3}{1-2x} = x^3(1 + 2x + (2x)^2 + \cdots) = x^3 + 2x^4 + 4x^5 + \cdots$$

Moreover, the term-by-term integration and differentiation theorems yield additional results in conjuction with the geometric series:

$$\tan^{-1}(x) = \int \frac{dx}{1+x^2} = \int \sum_{n=0}^{\infty} (-1)^n x^{2n} dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 + \cdots$$

$$\ln(1-x) = \int \frac{d}{dx} \ln(1-x) dx = \int \frac{-1}{1-x} dx = -\int \sum_{n=0}^{\infty} x^n dx = \sum_{n=0}^{\infty} \frac{-1}{n+1} x^{n+1}$$

Of course, these are just the basic building blocks. We also can twist things and make the student use algebra,

$$e^{x+2} = e^x e^2 = e^2(1 + x + \frac{1}{2}x^2 + \cdots)$$

or trigonmetric identities,

$$\sin(x) = \sin(x - 2 + 2) = \sin(x-2)\cos(2) + \cos(x-2)\sin(2)$$

$$\Rightarrow \quad \sin(x) = \cos(2) \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (x-2)^{2n+1} + \sin(2) \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} (x-2)^{2n}.$$

Feel free to peruse my most recent calculus II materials to see a host of similarly sneaky calculations.

7.2 solutions at an ordinary point

An **ordinary point** for a differential equation is simply a point at which an analytic solution exists. I'll explain more carefully how to discern the nature of a given ODE in the next section. In this section we make the unfounded assumption that a power series solution exists in each example.

Example 7.2.1. Problem: *find the first four nontrivial terms in a series solution centered at* $a = 0$ *for* $y' - y = 0$

Solution: *propose that* $y = c_o + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + c_5 x^5 \cdots$. *Differentiating,*

$$y' = c_1 + 2c_2 x + 3c_3 x^2 + 4c_4 x^3 + 5c_5 x^4 \cdots$$

We desire y be a solution, therefore:

$$y' - y = c_1 + 2c_2x + 3c_3x^2 + 4c_4x^3 + 5c_5x^4 \cdots - (c_o + c_1x + c_2x^2 + c_3x^3 + c_4x^4 + c_5x^5 \cdots) = 0.$$

Collect like terms:

$$c_1 - c_o + x(2c_2 - c_1) + x^2(3c_3 - c_2) + x^3(4c_4 - c_3) + x^4(5c_5 - c_4) + \cdots = 0$$

We find, by equating coefficients, that every coefficient on the l.h.s. of the expression above is zero thus:

$$c_1 = c_o, \quad c_2 = \frac{1}{2}c_1, \quad c_3 = \frac{1}{3}c_2, \quad c_4 = \frac{1}{4}c_3$$

Hence,

$$c_1 = c_o, \quad c_2 = \frac{1}{2}c_o, \quad c_3 = \frac{1}{3}\frac{1}{2}c_o, \quad c_4 = \frac{1}{4}\frac{1}{3}\frac{1}{2}c_o$$

Note that $2 = 2!, 3 \cdot 2 = 3!, 4 \cdot 3 \cdot 2 = 4!$ hence,

$$y = c_o + c_ox + \frac{1}{2}c_ox^2 + \frac{1}{3!}c_ox^3 + \frac{1}{4!}c_ox^4 + \cdots$$

Consequently, $\boxed{y = c_o(1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots)}$ *is the desired solution.*

Of course the example above is not surprising; $y' - y = 0$ has $\lambda - 1 = 0$ hence $y = c_oe^x$ is the solution. We just derived the first few terms in the power series expansion for e^x centered at $a = 0$.

Example 7.2.2. Problem: *find the complete series solution centered at $a = 0$ for $y'' + x^2y = 0$.*

Solution: *Suppose the solution is a power series and calculate,*

$$y = \sum_{k=0}^{\infty} c_kx^k, \quad y' = \sum_{k=0}^{\infty} kc_kx^{k-1}, \quad y'' = \sum_{k=0}^{\infty} k(k-1)c_kx^{k-2}$$

Of course, the summations can be taken from $k = 1$ for y' and $k = 2$ for y'' as the lower order terms vanish. Suppose $y'' + x^2y = 0$ to find:

$$\sum_{k=2}^{\infty} k(k-1)c_kx^{k-2} + x^2\sum_{k=0}^{\infty} c_kx^k = 0$$

Notice,

$$x^2\sum_{k=0}^{\infty} c_kx^k = \sum_{k=0}^{\infty} c_kx^2x^k = \sum_{k=0}^{\infty} c_kx^{k+2} = \sum_{j=2}^{\infty} c_{j-2}x^j$$

where in the last step we set $j = k + 2$ hence $k = 0$ gives $j = 2$. Likewise, consider:

$$\sum_{k=2}^{\infty} k(k-1)c_kx^{k-2} = \sum_{j=0}^{\infty}(j+2)(j+1)c_{j+2}x^j.$$

where we set $k - 2 = j$ hence $k = 2$ gives $j = 0$. Hence,

$$\sum_{j=0}^{\infty}(j+2)(j+1)c_{j+2}x^j + \sum_{j=2}^{\infty} c_{j-2}x^j = 0.$$

Sometimes we have to separate a few low order terms to clarify a pattern:

$$2c_2 + 6c_3x + \sum_{j=2}^{\infty}\big[(j+2)(j+1)c_{j+2} + c_{j-2}\big]x^j = 0$$

It follows that $c_2 = 0$ and $c_3 = 0$. Moreover, for $j = 2, 3, \dots$ we have the recursive rule:

$$c_{j+2} = \frac{-1}{(j+2)(j+1)}c_{j-2}$$

Let us study the relations above to find a pattern if possible,

$$c_4 = \frac{-1}{12}c_o, \quad c_5 = \frac{-1}{20}c_1, \quad c_6 = \frac{-1}{30}c_2, \quad c_7 = \frac{-1}{42}c_3, \quad c_8 = \frac{-1}{56}c_4, ...$$

Notice that $c_2 = 0$ clearly implies $c_{4k+2} = 0$ for $k \in \mathbb{N}$. Likewise, $c_3 = 0$ clearly implies $c_{4k+3} = 0$ for $k \in \mathbb{N}$. However, the coefficients $c_o, c_4, c_8, \dots$ are linked as are $c_1, c_5, c_9, \dots$. In particular,

$$c_{12} = \frac{-1}{(12)(11)}c_8 = \frac{-1}{(12)(11)}\cdot\frac{-1}{(8)(7)}c_4 = \frac{-1}{(12)(11)}\cdot\frac{-1}{(8)(7)}\cdot\frac{-1}{(4)(3)}c_o = c_{3(4)}$$

$$c_{16} = \frac{-1}{(16)(15)}\cdot\frac{-1}{(12)(11)}\cdot\frac{-1}{(8)(7)}\cdot\frac{-1}{(4)(3)}c_o = c_{4(4)}$$

We find,

$$c_{4k} = \frac{(-1)^k}{k!4^k(4k-1)(4k-5)\cdots 11\cdot 7\cdot 3}c_o$$

Next, study $c_1, c_5, c_9, ...$

$$c_9 = \frac{-1}{(9)(8)}c_5 = \frac{-1}{(9)(8)}\cdot\frac{-1}{(5)(4)}c_1 = c_{2(4)+1}$$

$$c_{13} = \frac{-1}{(13)(12)}\cdot\frac{-1}{(9)(8)}\cdot\frac{-1}{(5)(4)}c_1 = c_{3(4)+1}$$

We find,

$$c_{4k+1} = \frac{(-1)^k}{k!4^k(4k+1)(4k-3)\cdots 13\cdot 9\cdot 5}c_1$$

We find the solution has two coefficients c_o, c_1 as we ought to expect for the general solution to a second order ODE.

$$\boxed{y = c_o\sum_{k=0}^{\infty}\frac{(-1)^k}{k!4^k(4k-1)(4k-5)\cdots 11\cdot 7\cdot 3}x^{4k} + c_1\sum_{k=0}^{\infty}\frac{(-1)^k}{k!4^k(4k+1)(4k-3)\cdots 13\cdot 9\cdot 5}x^{4k+1}}$$

If we just want the the solution up to 11-th order in x then the following would have sufficed:

$$y = c_o(1 - \tfrac{1}{12}x^4 + \tfrac{1}{672}x^8 + \cdots) + c_1(x - \tfrac{1}{20}x^5 + \tfrac{1}{1440}x^9 + \cdots).$$

Remark 7.2.3.

The formulas we derived for c_{4k} and c_{4k+1} are what entitle me to claim the solution is the **complete** solution. It is not always possible to find nice formulas for the general term in the solution. Usually if no "nice" formula can be found you might just be asked to find the first 6 nontrvial terms since this typically gives 3 terms in each fundamental solution to a second order problem. We tend to focus on second order problems in this chapter, but most of the techniques here apply equally well to arbitrary order.

Example 7.2.4. Problem: *find the complete series solution centered at $a = 0$ for $y''+xy'+3y = 0$.*

Solution: *Suppose the solution is a power series and calculate,*

$$y = \sum_{k=0}^{\infty} c_k x^k, \quad y' = \sum_{k=0}^{\infty} k c_k x^{k-1}, \quad y'' = \sum_{k=0}^{\infty} k(k-1) c_k x^{k-2}$$

Suppose $y'' + xy' + 3y = 0$ to find:

$$\sum_{k=0}^{\infty} k(k-1) c_k x^{k-2} + x \sum_{k=0}^{\infty} k c_k x^{k-1} + 3 \sum_{k=0}^{\infty} c_k x^k = 0.$$

Hence, noting some terms vanish and $xx^{k-1} = x^k$:

$$\sum_{k=2}^{\infty} k(k-1) c_k x^{k-2} + \sum_{k=1}^{\infty} k c_k x^k + \sum_{k=0}^{\infty} 3 c_k x^k = 0$$

Let $k - 2 = j$ to relable $k(k-1)c_k x^{k-2} = (j+2)(j+1)c_{j+2} x^j$. It follows that:

$$\sum_{j=0}^{\infty} (j+2)(j+1) c_{j+2} x^j + \sum_{j=1}^{\infty} j c_j x^j + \sum_{j=0}^{\infty} 3 c_j x^j = 0$$

We can combine all three sums for $j \geq 1$ however the constant terms break the pattern so list them seperately,

$$2c_2 + 3c_o + \sum_{j=1}^{\infty} \Big[(j+2)(j+1) c_{j+2} + (3+j) c_j \Big] x^j = 0$$

Equating coefficients yields, for $j = 1, 2, 3 \dots$:

$$2c_2 + 3c_o = 0, \quad (j+2)(j+1)c_{j+2} + (3+j)c_j = 0 \quad \Rightarrow c_2 = \frac{-3}{2} c_o, \quad c_{j+2} = \frac{-(j+3)}{(j+2)(j+1)} c_j.$$

In this example the even and odd coefficients are linked. Let us study the recurrence relation above to find a general formula if possible.

$$\begin{aligned}
(j=1):& \quad c_3 = \frac{-4}{(3)(2)} c_1 = \frac{(-1)^1 (2^1)(2!)}{3!} c_1 \\
(j=3):& \quad c_5 = \frac{-6}{(5)(4)} c_3 = \frac{-6}{(5)(4)} \cdot \frac{-4}{(3)(2)} c_1 = \frac{(-1)^2 (2^2)(3!)}{5!} c_1 \\
(j=5):& \quad c_7 = \frac{-8}{(7)(6)} c_5 = \frac{-8}{(7)(6)} \cdot \frac{-6}{(5)(4)} \cdot \frac{-4}{(3)(2)} c_1 = \frac{(-1)^3 (2^3)(4!)}{7!} c_1 \\
(j=2k-1):& \quad c_{2k+1} = \frac{(-1)^k (2k+2)(2k)(2k-2)\cdots(6)(4)}{(2k+1)!} c_1.
\end{aligned}$$

Next, study the to even coefficients: we found $c_2 = \frac{-3}{2}c_o$

$$\begin{aligned}
(j=2): \quad & c_4 = \frac{-5}{(4)(3)}c_2 = \frac{-5}{(4)(3)} \cdot \frac{-3}{2}c_o \\
(j=4): \quad & c_6 = \frac{-7}{(6)(5)} \cdot \frac{-5}{(4)(3)} \cdot \frac{-3}{2}c_o \\
(j=6): \quad & c_8 = \frac{-9}{(8)(7)} \cdot \frac{-7}{(6)(5)} \cdot \frac{-5}{(4)(3)} \cdot \frac{-3}{2}c_o \\
(j=2k-2): \quad & c_{2k} = \frac{(-1)^k(2k+1)(2k-1)(2k-3)\cdots(7)(5)(3)}{(2k)!}c_o.
\end{aligned}$$

Therefore, the general solution is given by:

$$\boxed{y = c_o\sum_{k=0}^{\infty}\frac{(-1)^k(2k+1)(2k-1)\cdots(7)(5)(3)}{(2k)!}x^{2k} + c_1\sum_{k=0}^{\infty}\frac{(-1)^k(2k+2)(2k)\cdots(6)(4)}{(2k+1)!}x^{2k+1}}.$$

The first few terms in the solution are given by $y = c_o(1-\frac{3}{2}x^2+\frac{5}{8}x^4+\cdots)+c_1(x-\frac{2}{3}x^3+\frac{1}{5}x^5+\cdots)$.

Example 7.2.5. Problem: *find the first three nontrivial terms in the series solution centered at* $a=0$ *for* $y'' + \frac{1}{1-x}y' + e^x y = 0$. *Given that* $y(0)=0$ *and* $y'(0)=1$.

Solution: *Notice that* $\frac{1}{1-x} = 1+x+x^2+\cdots$ *and* $e^x = 1+x+\frac{1}{2}x^2+\cdots$ *hence:*

$$y'' + (1+x+x^2+\cdots)y' + (1+x+\tfrac{1}{2}x^2+\cdots)y = 0$$

Suppose $y = c_o+c_1x+c_2x^2+\cdots$ *hence* $y' = c_1+2c_2x+3c_3x^2+\cdots$ *and* $y'' = 2c_2+6c_3x+12c_4x^2+\cdots$. *Put these into the differential equation, keep only terms up to quadratic order,*

$$2c_2+6c_3x+12c_4x^2+(1+x+x^2)(c_1+2c_2x+3c_3x^2)+(1+x+\tfrac{1}{2}x^2)(c_o+c_1x+c_2x^2)+\cdots = 0$$

The coefficients of 1 *in the equation above are*

$$2c_2+c_1+c_o = 0$$

The coefficients of x *in the equation above are*

$$6c_3+c_1+2c_2+c_1+c_o = 0$$

The coefficients of x^2 *in the equation above are*

$$12c_4+c_1+2c_2+3c_3+\frac{1}{2}c_o+c_1+c_2 = 0$$

I find these problems very challenging when no additional information is given. However, we were given $y(0)=0$ *and* $y'(0)=1$ *hence*[1] $c_o = 0$ *whereas* $c_1 = 1$. *Thus* $c_2 = -1/2$ *and* $c_3 = \frac{-1}{6}(2c_2+2c_1) = \frac{-1}{6}$ *and* $c_4 = \frac{1}{12}(-2c_1-3c_2-3c_3) = \frac{1}{12}(1/2+3/2-2) = 0$ *hence*

$$\boxed{y = x - \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \cdots}.$$

[1] think about Taylor's theorem centered at zero

Remark 7.2.6.

When faced with a differential equation with variable coefficients we must expand the coefficient functions as power series when we seek a power series solution. Moreover, the center of the expansion ought to match the center of the desired solution. In this section we have only so far consider series centered at zero. Next we consider a nonzero center.

Example 7.2.7. **Problem:** *find the first few nontrivial terms in the series solution centered at* $a = 1$ *for* $y' = \frac{\sin(x)}{1-(x-1)^2}$.

Solution: *note that we can integrate to find an integral solution:* $y = \int \frac{\sin(x)\,dx}{1-(x-1)^2}$. *To derive the series solution we simply expand the integrand in powers of* $(x-1)$*. Note,*

$$\frac{1}{1-(x-1)^2} = 1 + (x-1)^2 + (x-1)^4 + (x-1)^6 + \cdots$$

On the other hand, to expand sine, we should use the adding angles formula on $\sin(x) = \sin(x-1+1)$ *to see*

$$\sin(x) = \cos(1)\sin(x-1) + \sin(1)\cos(x-1) = \sin(1) + \cos(1)(x-1) - \tfrac{\sin(1)}{2}(x-1)^2 + \cdots$$

Consider the product of the power series above, up to quadratic order we find:

$$\frac{\sin(x)}{1-(x-1)^2} = \sin(1) + \cos(1)(x-1) + \tfrac{\sin(1)}{2}(x-1)^2 + \cdots$$

Therefore, integrating term-by-term, we find

$$\boxed{y = c_1 + \sin(1)(x-1) + \tfrac{\cos(1)}{2}(x-1)^2 + \tfrac{\sin(1)}{6}(x-1)^3 + \cdots}.$$

Remark 7.2.8.

Taylor's formula $f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{6}f'''(a)(x-a)^3 + \cdots$ is one way we could compute the power series expansions for given functions, however, it is much faster to use algebra and known results when possible.

7.3 classification of singular points

We are primarily interested in real solutions to linear ODEs of first or second order in this chapter, however, the theory of singular points and the Frobenius method necessarily require us to consider singularities as having their residence in the complex plane. It would appear that our solutions are restrictions of complex solutions to the real axis in $\mathbb{C}$.

Definition 7.3.1. *singular points and ordinary points*

> We say x_o is a **ordinary point** of $y'' + Py' + Qy = 0$ iff P and Q are analytic at x_o. A point x_o is a **singular point** of $y'' + Py' + Qy = 0$ if x_o is not an ordinary point. A point x_o is a **regular singular point** of $y'' + Py' + Qy = 0$ if x_o is a singular point however $(x - x_o)P(x)$ and $(x - x_o)^2Q(x)$ are analytic at x_o.

In the definition above we mean to consider the functions $(x-x_o)P(x)$ and $(x-x_o)^2Q(x)$ with any removable discontinuities removed. For example, while $f(x) = \frac{1}{x}$ has $xf(x)$ undefined at $x = 0$, we still insist that $xf(x)$ is an analytic function at $x = 0$. Another example, technically the expression $\sin(x)/x$ is not defined at $x = 0$, but it is an analytic expression $1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^4 + \cdots$ which is defined at $x = 0$. To be more careful, we could insist that the limit as $x \to x_o$ of $(x - x_o)P(x)$ and $(x - x_o)^2Q(x)$ exist. That would just be a careful way of insisting that the only divergence faced by $(x-x_o)P(x)$ and $(x-x_o)^2Q(x)$ are simple holes in the graph a.k.a removable discontinuities.

In addition, the singular point x_o may be complex. This is of particular interest as we seek to determine the domain of solutions in the Frobenius method. I will illustrate by example:

Example 7.3.2. *For $b, c \in \mathbb{R}$, every point is an ordinary point for $y'' + by' + cy = 0$.*

Example 7.3.3. *Since e^x and $\cos(x)$ are analytic it follows that the differential equation $y'' + e^xy' + \cos(x)y = 0$ has no singular point. Every point is an ordinary point.*

Example 7.3.4. *Consider $(x^2+1)y''+y = 0$. We divide by x^2+1 and find $y'' + \frac{1}{x^2+1}y = 0$. Note:*

$$Q(x) = \frac{1}{x^2+1} = \frac{1}{(x+i)(x-i)}$$

It follows that every $x \in \mathbb{R}$ is an ordinary point and the only singular points are found at $x_o = \pm i$. It turns out that the existence of these imaginary singular points limits the largest open domain of a solution centered at the ordinary point $x_o = 0$ to $(-1, 1)$.

Example 7.3.5. *Consider $y'' + \frac{1}{x^2(x-1)}y' + \frac{1}{(x-1)^2(x^2+4x+5)}y = 0$. Consider,*

$$P(x) = \frac{1}{x^2(x-1)} \qquad \& \qquad Q(x) = \frac{1}{(x-1)^2(x-2+i)(x-2-i)}$$

Observe that,

$$xP(x) = \frac{x}{x^2(x-1)} = \frac{1}{x(x-1)}$$

therefore $xP(x)$ is not analytic at $x = 0$ hence $x = 0$ is a singular point which is not regular; this is also called an **irregular singular point**. *On the other hand, note:*

$$(x-1)P(x) = \frac{x-1}{x^2(x-1)} = \frac{1}{x^2} \qquad \& \qquad (x-1)^2Q(x) = \frac{(x-1)^2}{(x-1)^2(x^2+4x+5)} = \frac{1}{x^2+4x+5}$$

are both analytic at $x = 1$ hence $x = 1$ is a **regular singular point**. *Finally, note that the quadratic $x^2 + 4x + 5 = (x + 2 - i)(x + 2 + i)$ hence $x = -2 \pm i$ are singular points.*

It is true that $x = -2 \pm i$ are regular singular points, but this point does not interest us as we only seek solutions based at some real point.

Theorem 7.3.6. *ordinary points and frobenius' theorem*

A solution of $y'' + Py' + Qy = 0$ centered at an ordinary point x_o can be extended to an open disk in the complex plane which reaches the closest singularity. A solution of $y'' + Py' + Qy = 0$ based at a regular singular point x_o extends to an open interval with x_o at one edge and $x_o \pm R$ on the other edge where R is the distance to the next nearest singularity (besides x_o of course)

See pages 477 and 494 for corresponding theorems in Nagel, Saff and Snider. It is also important to note that the series technique and the full method of Frobenius will provide a fundamental solution set on the domains indicated by the theorem above.

Example 7.3.7. *Consider* $y'' + \frac{1}{x^2(x-1)}y' + \frac{1}{(x-1)^2(x^2+4x+5)}y = 0$. *Recall we found singular points* $x = 0, 1, -2+i, -2-i$. *The point* $x = 0$ *is an irregular singular point hence we have nothing much to say. On the other hand, if we consider solutions on* $(1, 1+R)$ *we can make* R *at most* $R = 1$ *the distance from* 1 *to* 0. *Likewise, we could find a solution on* $(0, 1)$ *which puts the regular singular point on the right edge. A solution* $\sum_{n=0}^{\infty} c_n(x+2)^n$ *centered at* $x = -2$ *will extend to the open interval* $(-3, -1)$ *at most since the singularities* $-2 \pm i$ *are one-unit away from* -2 *in the complex plane. On the other hand, if we consider a solution of the form* $\sum_{n=0}^{\infty} c_n(x+3)^n$ *which is centered at* $x = -3$ *then the singularities* $-2 \pm i$ *are distance* $\sqrt{2}$ *away and we can be confident the domain of the series solution will extend to at least the open interval* $(-3-\sqrt{2}, -3+\sqrt{2})$.

You might notice I was intentionally vague about the regular singular point solutions in the example above. We extend our series techniques to the case of a regular singular point in the next section.

7.4 frobenius method

We consider the problem $y'' + Py' + Qy = 0$ with a regular singular point x_o. We can study the case $x_o = 0$ without loss of generality since the subsitution $x = t - a$ moves the regular singular point to $t = a$. For example:

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + y = 0 \quad \Leftrightarrow \quad (t-a)^2\frac{d^2z}{dt^2} + (t-a)\frac{dz}{dt} + z = 0$$

Where $z(t) = y(x+a)$ and $y(x) = z(t-a)$. Therefore, we focus our efforts on the problem

$$\boxed{y'' + Py' + Qy = 0 \ \text{ a singular DEqn at } x = 0 \text{ with } xP(x), x^2Q(x) \text{ analytic at } x = 0}$$

Let us make some standard notation for the taylor expansions of $xP(x)$ and $x^2Q(x)$. Suppose

$$P(x) = \frac{P_o}{x} + P_1 + P_2x + \cdots \qquad \& \qquad Q(x) = \frac{Q_o}{x^2} + \frac{Q_1}{x} + Q_2 + Q_3x + \cdots$$

The extended Talyor series above are called **Laurent series**, they contain finitely many nontrivial reciprocal power terms. In the langauge of complex variables the pole $x = 0$ is removeable for P and Q where it is of order 1 and 2 respectively. Note we remove the singularity by multiplying by x and x^2:

$$xP(x) = P_o + P_1x + P_2x^2 + \cdots \qquad \& \qquad x^2Q(x) = Q_o + Q_1x + Q_2x^2 + Q_3x^3 + \cdots.$$

This must happen by the definition of a **regular singular point**.

Theorem 7.4.1. *frobenius solution at regular singular point*

There exists a number r and coefficients a_n such that $y'' + Py' + Qy = 0$ has solution

$$y = \sum_{n=0}^{\infty} a_n x^{n+r}.$$

See Rabenstein for greater detail as to why this solution exists. We can denote $y(r,x) = \sum_{n=0}^{\infty} a_n x^{n+r}$ if we wish to emphasize the dependence on r. Formally[2] it is clear that

$$y = \sum_{n=0}^{\infty} a_n x^{n+r} \ \ \& \ \ y' = \sum_{n=0}^{\infty} a_n(n+r)x^{n+r-1} \ \ \& \ \ y'' = \sum_{n=0}^{\infty} a_n(n+r)(n+r-1)x^{n+r-2}.$$

Notice that we make no assumption that $r = 0, 1, 2, \dots$ hence $y(r,x)$ is not necessarily a power series. The frobenius solution is more general than a simple power series. Let us continue to plug in the formulas for y, y', y'' into $x^2y'' + x^2Py' + x^2Qy = 0$:

$$\begin{aligned} 0 = x^2 &\sum_{n=0}^{\infty} a_n(n+r)(n+r-1)x^{n+r-2} \\ &+ \Big(P_o + P_1x + xP_2x^3 + \cdots\Big)x\sum_{n=0}^{\infty} a_n(n+r)x^{n+r-1} \\ &+ \Big(Q_o + Q_1x + Q_2x^2 + Q_3x^3 + \cdots\Big)\sum_{n=0}^{\infty} a_n x^{n+r} \end{aligned}$$

[2]formal in the sense that we ignore questions of convergence

Hence, (call this $\star$ for future reference)

$$0 = \sum_{n=0}^{\infty} a_n(n+r)(n+r-1)x^{n+r} + \Big(P_o + P_1 x + xP_2 x^3 + \cdots\Big)\sum_{n=0}^{\infty} a_n(n+r)x^{n+r} + \Big(Q_o + Q_1 x + Q_2 x^2 + Q_3 x^3 + \cdots\Big)\sum_{n=0}^{\infty} a_n x^{n+r}$$

You can prove that $\{x^r, x^{r+1}, x^{r+2}, \dots\}$ is a linearly independent set of functions on appropriate intervals. Therefore, $y(r,x)$ is a solution iff we make each coefficient vanish in the equation above. We begin by examining the $n=0$ terms which are the coefficient of x^r:

$$a_o(0+r)(0+r-1) + P_o a_o(0+r) + Q_o a_o = 0$$

This gives no condition on a_o, but we see that r must be chosen such that

$$\boxed{r(r-1) + rP_o + Q_o = 0} \qquad \text{the } \textbf{indicial} \text{ equation}$$

We find that we must begin the Frobenius problem by solving this equation. We are not free to just use any r, a particular pair of choices will be dictated from the zeroth coefficients of the xP and x^2Q Taylor expansions. Keeping in mind that r is not free, let us go on to describe the next set of equations from the coefficient of x^{r+1} of $\star$ $(n=1)$,

$$a_1(1+r)r + (1+r)P_o a_1 + rP_1 a_o + Q_o a_1 + Q_1 a_o = 0$$

The equation above links a_o to a_1. Next, for x^{r+2} in $\star$ we need

$$a_2(2+r)(r+1) + (2+r)P_o a_2 + (1+r)P_1 a_1 + rP_2 a_o + Q_o a_2 + Q_1 a_1 + Q_o a_2 = 0$$

The equation above links a_2 to a_1 and a_o. In practice, for a given problem, the recurrence relations which define a_k are best derived directly from $\star$. I merely wish to indicate the general pattern[3] with the remarks above.

Example 7.4.2. Problem: *solve* $3xy'' + y' - y = 0$.

Solution: *Observe that* $x_o = 0$ *is a regular singular point. Calculate,*

$$y = \sum_{n=0}^{\infty} a_n x^{n+r} = \sum_{n=1}^{\infty} a_{n-1} x^{n+r-1}$$

and

$$y' = \sum_{n=0}^{\infty} a_n(n+r)x^{n+r-1} \qquad \& \qquad 3xy'' = \sum_{n=0}^{\infty} 3a_n(n+r)(n+r-1)x^{n+r-1}$$

[3] if one wishes to gain a deeper calculational dexterity with this method I highly reccommend the sections in Rabenstein, he has a few techniques which are superior to the clumsy calculations I perform here

Therefore, $3xy'' + y' - y = 0$ yields

$$a_o[3r(r-1)+r]x^{r-1} + \sum_{n=1}^{\infty}\Big(3a_n(n+r)(n+r-1) + a_n(n+r) - a_{n-1}\Big)x^{n+r-1} = 0$$

Hence, for $n = 1, 2, 3, \dots$ we find:

$$3r(r-1)+r = 0 \qquad \& \qquad \boxed{a_n = \frac{a_{n-1}}{(n+r)(3n+3r-2)}}. \ \star$$

The indicial equation $3r(r-1)+r = 3r^2-2r = r(3r-2) = 0$ gives $r_1 = 2/3$ and $r_2 = 0$. Suppose $r_1 = 2/3$ and work out the recurrence relation $\star$ in this context: $a_n = \frac{a_{n-1}}{n(3n+2)}$ thus:

$$a_1 = \frac{a_o}{5}, \quad a_2 = \frac{a_1}{8\cdot 2} = \frac{a_o}{8\cdot 5\cdot 2}, \quad a_3 = \frac{a_2}{11\cdot 3} = \frac{a_o}{11\cdot 8\cdot 5\cdot 3\cdot 2}$$

$$a_4 = \frac{a_3}{14\cdot 4} = \frac{a_o}{14\cdot 11\cdot 8\cdot 5\cdot 4\cdot 3\cdot 2} \quad \Rightarrow \quad a_n = \frac{a_o}{5\cdot 8\cdot 11\cdot 14\cdots(3n+2)n!} \quad (n = 1, 2, \dots)$$

Therefore, $y(2/3, x) = a_o\big(x^{2/3} + \sum_{n=1}^{\infty}\frac{x^{n+2/3}}{5\cdot 8\cdot 11\cdot 14\cdots(3n+2)n!}\big)$ is a solution. Next, work out the recurrence relation $\star$ in the $r_2 = 0$ case: $a_n = \frac{a_{n-1}}{n(3n-2)}$ thus:

$$a_1 = \frac{a_o}{1}, \quad a_2 = \frac{a_1}{2\cdot 4} = \frac{a_o}{2\cdot 4}, \quad a_3 = \frac{a_2}{3\cdot 7} = \frac{a_o}{7\cdot 4\cdot 3\cdot 2}$$

$$a_4 = \frac{a_3}{4\cdot 10} = \frac{a_o}{10\cdot 7\cdot 4\cdot 4\cdot 3\cdot 2} \quad \Rightarrow \quad a_n = \frac{a_o}{4\cdot 7\cdot 10\cdots(3n-2)n!} \quad (n = 2, 3, \dots)$$

Consequently, $y(0, x) = a_o\big(1 + x + \sum_{n=2}^{\infty}\frac{x^n}{4\cdot 7\cdot 10\cdots(3n-2)n!}\big)$. We find the general solution

$$\boxed{y = c_1\left(x^{2/3} + \sum_{n=1}^{\infty}\frac{x^{n+2/3}}{5\cdot 8\cdot 11\cdot 14\cdots(3n+2)n!}\right) + c_2\left(1 + x + \sum_{n=2}^{\infty}\frac{x^n}{4\cdot 7\cdot 10\cdots(3n-2)n!}\right)}.$$

Remark 7.4.3.

Before we try another proper example I let us apply the method of Frobenius to a Cauchy Euler problem. The Cauchy Euler problem $x^2y'' + Pxy' + Qy = 0$ has $P_o = P$ and $Q_o = Q$. Moreover, the characteristic equation $r(r-1) + rP_o + Q_o = 0$ is the indicial equation. In other words, the regular singular point problem is a generalization of the Cauchy Euler problem. In view of this you can see our discussion thus far is missing a couple cases: (1.) the repeated root case needs a natural log, (2.) the complex case needs the usual technique. It turns out there is another complication. When r_1, r_2 are the exponents with $Re(r_1) > Re(r_2)$ and $r_1 - r_2$ is a positive integer we sometimes need a natural log term.

Example 7.4.4. Problem: *solve $x^2y'' + 3xy' + y = 0$.*

Solution: *Observe that $y'' + \frac{3}{x}y' + \frac{1}{x^2}y = 0$ has regular singular point $x_o = 0$ and $P_o = 3$ whereas $Q_o = 1$. The indicial equation $r(r-1) + 3r + 1 = r^2 + 2r + 1 = (r+1)^2 = 0$ gives $r_1 = r_2 = -1$. Suppose $y = y(-1, x) = \sum_{n=0}^{\infty} a_n x^{n-1}$. Plugging $y(-1, x)$ into $x^2y'' + 3xy' + y = 0$ yields:*

$$\sum_{n=0}^{\infty}(n-1)(n-2)a_n x^{n-1} + \sum_{n=0}^{\infty} 3(n-1)a_n x^{n-1} + \sum_{n=0}^{\infty} a_n x^{n-1} = 0$$

Collecting like powers is simple for the expression above, we find:

$$\sum_{n=0}^{\infty}\Big((n-1)(n-2)a_n + 3(n-1)a_n + a_n\Big)x^{n-1} = 0$$

Hence $[(n-1)(n-2) + 3(n-1) + 1]a_n = 0$ for $n = 0, 1, 2, \dots$. Put $n = 0$ to obtain $0a_o = 0$ hence no condition for a_o is found. In contrast, for $n \geq 1$ the condition yields $a_n = 0$. Thus $y(-1, x) = a_o x^{-1}$. Of course, you should have expected this from the outset! This is a Cauchy Euler problem, we expect the general solution $y = c_1 \frac{1}{x} + c_2 \frac{ln(x)}{x}$.

We examine a solution with imaginary exponents.

Example 7.4.5. Problem: *solve $x^2y'' + xy' + (4-x)y = 0$.*

Solution: *Observe that $x_o = 0$ is a regular singular point. Calculate, if $y = \sum_{n=0}^{\infty} a_n x^{n+r}$ then*

$$(4-x)y = \sum_{n=0}^{\infty} 4a_n x^{n+r} - \sum_{n=0}^{\infty} a_n x^{n+r+1} = \sum_{n=0}^{\infty} 4a_n x^{n+r} - \sum_{j=1}^{\infty} a_{j-1} x^{j+r}$$

and

$$xy' = \sum_{n=0}^{\infty} a_n(n+r)x^{n+r} \qquad \& \qquad x^2y'' = \sum_{n=0}^{\infty} a_n(n+r)(n+r-1)x^{n+r}$$

Therefore, $x^2y'' + xy' + (4-x)y = 0$ yields

$$a_o[r(r-1) + r + 4]x^r + \sum_{n=1}^{\infty}\Big(a_n(n+r)(n+r-1) + a_n(n+r) + 4a_n - a_{n-1}\Big)x^{n+r}$$

Hence, for $n = 1, 2, 3, \dots$ we find:

$$r^2 + 4 = 0 \qquad \& \qquad \boxed{a_n = \frac{a_{n-1}}{(n+r)^2 + 4}}. \quad \star$$

The indicial equation $r^2 + 4 = 0$ gives $r_1 = 2i$ and $r_2 = -2i$. We study $\star$ in a few cases. Let me begin by choosing $r = 2i$. Let's reformulate $\star$ into a cartesian form:

$$a_n = \frac{a_{n-1}}{(n+2i)^2 + 4} = \frac{a_{n-1}}{n^2 + 4ni - 4 + 4} = \frac{a_{n-1}}{n^2 + 4ni} \cdot \frac{n^2 - 4ni}{n^2 - 4ni} = \frac{a_{n-1}(n^2 - 4ni)}{n^2(n^2 + 16)} \quad \star^2$$

Consider then, by $\star^2$

$$a_1 = \frac{a_o(1-4i)}{17}, \qquad a_2 = \frac{a_1(4-8i)}{4(4+16)} = \frac{a_o(1-4i)}{17} \cdot \frac{4-8i}{80} = \frac{-a_o(28+24i)}{(17)(80)}$$

Consequently, we find the complex solution are

$$\begin{aligned} y &= a_o\left(x^{2i} + \frac{1-4i}{17}x^{1+2i} - \frac{28+12i}{(17)(80)}x^{2+2i} + \cdots\right) \\ &= a_o x^{2i}\Bigg(\underbrace{1 + \frac{1}{17}x - \frac{28}{(17)(80)}x^2 + \cdots}_{a(x)} + i\underbrace{\left[\frac{-4}{17}x - \frac{12}{(17)(80)}x^2 + \cdots\right]}_{b(x)}\Bigg) \end{aligned}$$

Recall, for $x > 0$ we defined $x^{n+2i} = x^n[\cos(2\ln(x)) + i\sin(2\ln(x))]$. Therefore,

$$y = a_o\Big[\cos(2\ln(x))a(x) - \sin(2\ln(x))b(x)\Big] + ia_o\Big[\sin(2\ln(x))a(x) + \cos(2\ln(x))b(x)\Big]$$

forms the general complex solution. Set $a_o = 1$ to select the real fundmental solutions $y_1 = Re(y)$ and $y_2 = Im(y)$. The general real solution is $y = c_1y_1 + c_2y_2$. In particular,

$$\boxed{y = c_1\Big[\cos(2\ln(x))a(x) - \sin(2\ln(x))b(x)\Big] + c_2\Big[\sin(2\ln(x))a(x) + \cos(2\ln(x))b(x)\Big]}$$

We have made manifest the first few terms in a and b, it should be clear how to find higher order terms through additional iteration on $\star^2$. The proof that these series converge can be found in more advanced sources (often Ince is cited by standard texts).

Remark 7.4.6.

> The calculation that follows differs from our initial example in one main aspect. I put in the exponents before I look for the recurrence relation. It turns out that the method of Example 7.4.2 is far more efficient a method of calculation. I leave this slightly clumsy calculation to show you the difference. You should use the approach of Example 7.4.2 for brevity's sake..

Example 7.4.7. Problem: *solve $xy'' + (3 + x^2)y' + 2xy = 0$.*

Solution: *Observe $y'' + (3/x + x)y' + 2y = 0$ thus identify that $P_o = 3$ whereas $Q_o = 0$. The indicial equation $r(r-1) + 3r = 0$ yields $r(r+2) = 0$ thus the* **exponents** *are $r_1 = 0, r_2 = -2$. In order to find the coefficients of $y(0, x) = y = \sum_{n=0}^{\infty} a_n x^n$ we must plug this into $xy'' + 3y' + x^2y' + 2xy = 0$,*

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-1} + \sum_{n=1}^{\infty} 3na_n x^{n-1} + \sum_{n=1}^{\infty} na_n x^{n+1} + \sum_{n=0}^{\infty} 2a_n x^{n+1} = 0$$

Examine these summations and note that x^1, x^0, x^2, x^1 are the lowest order terms respectively from left to right. To combine these we will need to start with x^2-terms.

$$\begin{aligned} 0 &= 2a_2x + 3a_1 + 6a_2x + 2a_ox \\ &\quad + \sum_{n=3}^{\infty} n(n-1)a_n x^{n-1} + \sum_{n=3}^{\infty} 3na_n x^{n-1} + \sum_{n=1}^{\infty} na_n x^{n+1} + \sum_{n=1}^{\infty} 2a_n x^{n+1} = 0 \end{aligned}$$

Let $j = n - 1$ for the first two sums and let $j = n + 1$ for the next two sums.

$$0 = 3a_1 + (2a_2 + 6a_2 + 2a_o)x + \sum_{j=2}^{\infty}\Big((j+1)ja_{j+1} + 3(j+1)a_{j+1} + (j-1)a_{j-1} + 2a_{j-1}\Big)x^j$$

Collecting like terms we find:

$$0 = 3a_1 + (8a_2 + 2a_o)x + \sum_{j=2}^{\infty}\Big((j+3)(j+1)a_{j+1} + (j+1)a_{j-1}\Big)x^j.$$

Each power's coefficient must separately vanish, therefore:

$$a_1 = 0, \qquad a_2 = -\frac{1}{4}a_o, \qquad a_{j+1} = \frac{-1}{j+3}a_{j-1}, \qquad \Rightarrow \qquad a_n = \frac{-1}{n+2}a_{n-2} \ \ \textit{for } n \geq 2.$$

It follows that $a_{2k+1} = 0$ for $k = 0, 1, 2, 3, \ldots$. However, the even coefficients are determined by the recurrence relation given above.

$$a_2 = \frac{-1}{4}a_o$$

$$a_4 = \frac{-1}{6}a_2 = \frac{-1}{6}\cdot\frac{-1}{4}a_o$$

$$a_6 = \frac{-1}{8}a_4 = \frac{-1}{8}\cdot\frac{-1}{6}\cdot\frac{-1}{4}a_o$$

$$a_{2k} = \frac{-1}{2k+2}\cdot\frac{-1}{2k}\cdots\frac{-1}{6}\cdot\frac{-1}{4}a_o = \frac{(-1)^k}{2^k(k+1)!}a_o$$

Therefore, we find the solution:

$$\boxed{y(0,x) = a_o\sum_{k=0}^{\infty}\frac{(-1)^k}{2^k(k+1)!}x^{2k}}$$

We know from the theory we discussed in previous chapters the general solution should have the form $y = c_1y_1 + c_2y_2$ where $\{y_1, y_2\}$ is the fundamental solution set. We have found half of the solution at this point; identify $y_1 = y(0,x)$. In contrast to the series method, we found just one of the fundamental solutions.

To find y_2 we must turn our attention to the second solution of the indicial equation $r_2 = -2$. We find the coefficients of $y(-2,x) = y = \sum_{n=0}^{\infty} a_nx^{n-2}$ by plugging it into $xy'' + 3y' + x^2y' + 2xy = 0$,

$$\sum_{n=0}^{\infty}(n-2)(n-3)a_nx^{n-3} + \sum_{n=0}^{\infty}3(n-2)a_nx^{n-3} + \sum_{n=0}^{\infty}(n-2)a_nx^{n-1} + \sum_{n=0}^{\infty}2a_nx^{n-1} = 0$$

1. *x^{-3} has coefficient $(-2)(-3)a_o + 3(-2)a_o = 0$ (no condtion found)*
2. *x^{-2} has coefficient $(1-2)(1-3)a_1 + 3(1-2)a_1 = -a_1$ hence $a_1 = 0$*
3. *x^{-1} has coefficient $(2-2)(2-3)a_2 + 3(2-2)a_2 + (0-2)a_o + 2a_o = 0$ (no condtion found)*
4. *x^0 has coefficient $(3-2)(3-3)a_3 + 3(3-2)a_3 + (1-2)a_1 + 2a_1 = 3a_3$. Thus $a_3 = 0$*

5. *x^1 has coefficient $(4-2)(4-3)a_4+3(4-2)a_4+(2-2)a_2+2a_2 = 8a_4+2a_2$. Thus $a_4 = \frac{-1}{4}a_2$.*

6. *x^2 has coefficient $(5-2)(5-3)a_5+3(5-2)a_5+(3-2)a_3+2a_3 = 15a_5+3a_3$. Thus $a_5 = \frac{-1}{5}a_3$. We find $a_{2k-1} = 0$ for all $k \in \mathbb{N}$.*

7. *x^3 has coefficient $(6-2)(6-3)a_6+3(6-2)a_6+(4-2)a_4+2a_4 = 24a_6+4a_4$. Thus $a_6 = \frac{-1}{6}a_4$.*

This pattern should be recongnized from earlier in this problem. For $a_2, a_4, a_6, \dots$ we find terms

$$a_2 - a_2\tfrac{1}{4}x^2 + a_2\tfrac{1}{4}\cdot\tfrac{1}{6}x^4 + \cdots = a_2(1 - \tfrac{1}{2^1 2!}x^2 + \tfrac{1}{2^2 3!}x^4 + \cdots)$$

recognize this is simply a relabeled version of $y(0,x)$ hence we may set $a_2 = 0$ without loss of generality in the general solution. This means only a_o remains nontrivial. Thus,

$$\boxed{y(-2,x) = a_o x^{-2}}$$

The general solution follows,

$$\boxed{y = c_1 \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k (k+1)!} x^{2k} + \frac{c_2}{x^2}}$$

Remark 7.4.8.

In the example above the exponents $r_1 = 0$ and $r_2 = -2$ have $r_1 - r_2 = 2$. It turns out that generally the solution $y(r_2, x)$ will not be a solution. A modification involving a logarithm is needed sometimes (but not in the example above!).

7.4.1 the repeated root technique

In the case that a characteristic root is repeated we have seen the need for special techniques to derive a second LI solution. I present a new idea, yet another way to search for such double-root solutions. Begin by observing that the double root solutions are connected to the first solution by differentiation of the characteristic value:

$$\frac{\partial}{\partial r}e^{rx} = xe^{rx}, \qquad \& \qquad \frac{\partial}{\partial r}x^r = \ln(x)x^r.$$

Rabenstein gives a formal derivation of why $\frac{\partial x^r}{\partial r}\Big|_{r=r_1}$ solves a Cauchy Euler problem with repeated root r_1. I'll examine the corresponding argument for the repeated root case $(D^2 - 2\lambda_1 D + \lambda_1^2)[y] = L[y] = 0$. Suppose $y(\lambda, x) = e^{\lambda x}$. Note that:

$$L[e^{\lambda x}] = (\lambda^2 - 2\lambda_1\lambda + \lambda_1^2)e^{\lambda x} = (\lambda - \lambda_1)^2 e^{\lambda x}$$

Obviously $y_1 = y(\lambda_1, x)$ solves $L[y] = 0$. Consider $y_2 = \frac{\partial}{\partial\lambda}y(\lambda,x)\big|_{\lambda=\lambda_1}$

$$\begin{aligned} L[y_2] = L\left[\frac{\partial}{\partial\lambda}y(\lambda,x)\bigg|_{\lambda=\lambda_1}\right] &= \frac{\partial}{\partial\lambda}\Big[L[y(\lambda,x)]\Big]\bigg|_{\lambda=\lambda_1} \\ &= \frac{\partial}{\partial\lambda}\Big[(\lambda-\lambda_1)^2 e^{\lambda x}\Big]\bigg|_{\lambda=\lambda_1} \\ &= \Big[2(\lambda-\lambda_1)e^{\lambda x} + (\lambda-\lambda_1)^2 x e^{\lambda x}\Big]\bigg|_{\lambda=\lambda_1} \\ &= 0. \end{aligned}$$

Suppose that we face $x^2y'' + Px^2y' + x^2Q = 0$ which has an indicial equation with repeated root r_1. Suppose[4] $y(r,x) = x^r \sum_{n=0}^{\infty} a_n(r)x^n$ is a solution $x^2y'' + Px^2y' + x^2Q = 0$ when we set $r = r_1$. It can be shown[5] that $y_2 = \frac{\partial y(r,x)}{\partial r}\big|_{r=r_1}$ solves $x^2y'' + Px^2y' + x^2Q = 0$. Consider,

$$\frac{\partial y(r,x)}{\partial r} = \frac{\partial}{\partial r}\left[x^r \sum_{n=0}^{\infty} a_n(r)x^n\right] = \ln(x)x^r \sum_{n=0}^{\infty} a_n(r)x^n + x^r \sum_{n=0}^{\infty} a_n'(r)x^n$$

Setting $r = r_1$ and denoting $y_1(x) = y(r_1,x) = x^{r_1} \sum_{n=0}^{\infty} a_n(r_1)x^n$ we find the second solution

$$\boxed{y_2(x) = \ln(x)y_1(x) + x^{r_1} \sum_{n=0}^{\infty} a_n'(r_1)x^n.}$$

Compare this result to Theorem 7 of section 8.7 in Nagel Saff and Snider to appreciate the beauty of this formula. If we calculate the first solution then we find the second by a little differentiation and evaluation at r_1.

Example 7.4.9. *include example showing differentiation of $a_n(r)$ (to be given in lecture most likely)*

Turn now to the case $x^2y'' + Px^2y' + x^2Q = 0$ has exponents r_1, r_2 such that $r_1 - r_2 = N \in \mathbb{N}$. Following Rabenstein once more I examine the general form of the recurrence relation that formulates the coefficients in the Frobenius solution. We will find that for r_1 the coefficients exist, however for r_2 there exist P, Q such that the recurrence relation is insolvable. We seek to understand these features.

Remark 7.4.10.

Sorry these notes are incomplete. I will likely add comments based on Rabenstein in lecture, his treatment of Frobenius was more generous than most texts at this level. In any event, you should remember these notes are a work in progress and you are welcome to ask questions about things which are not clear.

7.5 practice problems

PP 183 Find the first three nonzero terms in the power series solutions of

$$\frac{dy}{dx} = x^2 + y^2$$

given $y(0) = 1$.

PP 184 Find the first three nonzero terms in the power series solutions of

$$\frac{dy}{dx} = \sin y + e^x$$

given $y(0) = 0$.

[4] I write the x^r in front and emphasize the r-dependence of the a_n coefficients as these are crucial to what follows, if you examine the previous calculations you will discover that a_n does dependend on the choice of exponent

[5] see Rabenstein page 120

PP 185 Find the first three nonzero terms in the power series solutions of

$$x'' + tx = 0$$

given $x(0) = 1$ and $x'(0) = 0$.

PP 186 **Duffing's Equation.** A nonlinear spring with periodic forcing is described by

$$y'' + ky + ry^3 = A\cos\omega t.$$

If we set $k = r = A = 1$ and $\omega = 10$ then find the first three nonzero terms in the Taylor polynomial approximations tot he solution with $y(0) = 0$ and $y'(0) = 1$.

PP 187 Express the power series $\sum_{n=1}^{\infty} na_n x^{n-1}$ as a power series with generic term x^k. That is, find k_o and c_k for which $\sum_{n=1}^{\infty} na_n x^{n-1} = \sum_{k=k_o}^{\infty} c_k x^k$.

PP 188 Express the power series $\sum_{n=1}^{\infty} a_n x^{n+1}$ as a power series with generic term x^k. That is, find k_o and c_k for which $\sum_{n=1}^{\infty} a_n x^{n+1} = \sum_{k=k_o}^{\infty} c_k x^k$.

PP 189 Find the Taylor series for $f(x) = \dfrac{1+x}{1-x}$ about $x_0 = 0$.

PP 190 Find the singular points of the differential equation $(x+1)y'' - x^2y' + 3y = 0$.

PP 191 Find the singular points of the differential equation $(t^2 - t - 2)x'' - (t+1)x' - (t-2)x = 0$.

PP 192 Find the first four nonzero terms in the power series solution about $x = 0$ for:

$$z'' - x^2 z = 0.$$

PP 193 Find the first four nonzero terms in the power series solution about $x = 0$ for:

$$y'' + (x-1)y' + y = 0.$$

PP 194 Find the complete power series solution (including a formula for the general coefficient) about $x = 0$ for:

$$y' - 2xy = 0.$$

PP 195 Find the complete power series solution (including a formula for the general coefficient) about $x = 0$ for:

$$y'' - xy' + 4y = 0.$$

PP 196 Find the complete power series solution (including a formula for the general coefficient) about $x = 0$ for:

$$z'' - x^2 z' - xz = 0.$$

PP 197 Find the minimum value for the radius of covergence of a power series solution about x_0

$$(1+x+x^2)y'' - 3y = 0, \qquad x_0 = 1.$$

PP 198 Find the minimum value for the radius of covergence of a power series solution about x_0

$$y'' - (\tan x)y' + y = 0 = 0, \qquad x_0 = 0.$$

PP 199 Find the first four nonzero terms in the power series solution about $x = 0$ for:

$$x' + (\sin t)x = 0, \qquad x(0) = 1.$$

PP 200 Find the first four nonzero terms in the power series solution about $x = 0$ for:

$$y'' - e^{2x}y' + (\cos x)y = 0, \qquad y(0) = -1,\ y'(0) = 1.$$

PP 201 Find the first four nonzero terms in the power series solution about $x = 0$ for:

$$z'' + xz' + z = x^2 + 2x + 1.$$

PP 202 Find the first four nonzero terms in the power series solution about $x = 0$ for:

$$(1+x^2)y'' - xy' + y = e^{-x}.$$

PP 203 If $\sum_{n=0}^{\infty} B_n x^n = \sum_{n=2}^{\infty} n(n-1)c_n x^{n-2} + \sum_{n=0}^{\infty} c_n x^{n+2}$ then find the formula for B_n in terms of c_n. You will need to break into cases, B_0, B_1 verse B_n for $n \geq 2$.

PP 204 Find the minimum radius of convergence about $x = 0$ for the solution of

$$(x^2 - 2x + 10)y'' + xy' - 4y = 0.$$

PP 205 Solve $y'' + (x+1)y' - y = 0$ up to 4-th order. Center the solution at zero.

PP 206 Find the first three nontrivial terms in the power series solution centered at zero of the differential equation $(x^2+1)y'' + 2xy' = 0$ with $y(0) = 0$ and $y'(0) = 1$.

PP 207 Is $x = 0$ an ordinary point of $y'' + 5xy' + \sqrt{x}y = 0$?

PP 208 Find all singularties of the following differential equations, or state no singularities:

(a.) $y'' + xy' + 3y = 0,$

(b.) $(x^2 - 3x_2)y'' + \sqrt{x}y' + x^2y = 0$

(c.) $(1-x^2)y'' - 2xy' + n(n+1)y = 0$

(d.) $(x^2 - x)y'' + x^2y' - 3xy = 0$

(e.) $e^x - 1)y'' + xy = 0$

(f.) $x(x^2+2x+2)y'' + (x^2+1)y' + 3y = 0$

PP 209 Find the complete Frobenius solution of

$$x^2 y'' + x\left(x - \frac{1}{2}\right)y' + \frac{1}{2}y = 0$$

(this one has real exponents $r = 1$ and $r = 1/2$)

PP 210 Find the Frobenius solution near $x = 0$ for $x > 0$ up to order x^2 for

$$x^2 y'' + \sin(x)y' - \cos(x)y = 0.$$

PP 211 Solve $x^3 y'' - x^2 y' - y = 0$ for $x >> 0$ by making the substitution $z = 1/x$ and solving the resulting differential equation in z about the regular singular point $z = 0$. Find the first four nonzero terms in the series expansion about ∞ (once upon a time this was Problem 41 in §8.6 of Nagle, Saff and Snider, 5th edition)

PP 212 Find the complete (summation-notation) power series solution of the following integral:

$$\int x^6 \sin(x^2)\, dx$$

PP 213 Find the first **TWO** nontrivial terms in a power series solution of $e^x y'' + xy' + y = 0$ given that $y(0) = 1$ and $y'(0) = 2$.

PP 214 Find the singularities of $x(x^2 + 2x + 2)y'' + (x^2 + 1)y' + 3y = 0$ and determine the largest open interval of convergence for a solution of the form $y = \sum_{n=0} a_n (x + 2)^n$.
Think. Do not try to solve this, I'm asking you about the interval of convergence, I'm ***not*** *asking for what a_n are in particular*

PP 215 Find the complete power series solution of $y'' - 9x^2 y = 0$ given that $y(0) = 1$ and $y'(0) = 0$ by explicit substitution of a series solution into the given differential equation.

PP 216 Suppose $y'' + \frac{x}{(x-2)(x^2-6x+10)}y' + \left(\frac{1}{(x+3)^3} + \frac{1}{x^2}\right)y = 0$.

(a) find all singular points

(b) classify each real singular point as either regular or irregular (not regular)

(c) plot the singularities in a complex plane

(d) find the largest possible open and real domain of the solution

$$y = \sum_{n=0}^{\infty} a_n (x - 0.5)^2$$

(e) find the largest possible open and real domain of the solution

$$y = \sum_{n=0}^{\infty} a_n (x - 4)^2$$

PP 217 Suppose $y(0) = 1$ and $y'(0) = 2$. Find the solution up to order 5 in x for the differential equation

$$y'' + (x^2 - 1)\cos(x)y' + \sinh(3x)y = 0.$$

PP 218 Find the complete power series solution centered at zero for $\frac{dy}{dx} - 2xy = 0$.

PP 219 Find the first two nontrivial terms in the frobenius expansions for the fundamental solutions y_1 and y_2 of

$$3xy'' + (2-x)y' - y$$

PP 220 Find the complete power series solution of $y'' + x^2y' + 2xy = 0$ about the ordinary point $x = 0$. Your answer should include nice formulas for arbitrary coefficients in each of the fundamental solutions. You need to both set-up and solve the reccurrence relations as best you can.

PP 221 Find the first four nonzero terms in the power series solution about zero for the initial value problem $y'' + \sin(x)y' + (x-1)y = 0$ with $y(0) = 1$ and $y'(0) = 0$.

PP 222 Find the complete Frobenius solution of

$$x^2y'' + x(x - \tfrac{1}{2})y' + \tfrac{1}{2}y = 0.$$

(it turns out this one has real exponents)

PP 223 Solve $x^3y'' - x^2y' - y = 0$ for $x >> 0$ by making the subsitution $z = 1/x$ and solving the resulting differential equation in z about the regular singular point $z = 0$. Find the first four nonzero terms in the series expansion about infinity.

PP 224 Consider $y'' + e^xy' + \sin(3x)y = 0$. Find the first 3 nontrivial terms in a series solution centered about $x = 0$ given that $y(0) = 1$ and $y'(0) = 6$.

PP 225 Find the complete power series solution of $y'' + 6x^2y = 0$ centered at $x = 0$.

PP 226 Suppose we define $e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n$. Show that $e^{i\theta} = \cos(\theta) + i\sin(\theta)$.

PP 227 Suppose $\sum_{k=0}^{\infty}(a_{2k}x^{2k} + b_{2k+1}x^{2k+1} = e^x + \cos(x+2)$. Find explicit formulas for a_{2k} and b_{2k+1} via Σ-notation algebra.

PP 228 Find a power series solution to the integrals below:

(a.) $\displaystyle\int \frac{x^3 + x^6}{1 - x^3}\, dx$

(b.) $\displaystyle\int x^8 e^{x^3+2}\, dx$

PP 229 Calculate the 42^{nd}-derivative of $x^2\cos(x)$ at $x = 1$. (use power series techniques)

PP 230 Find the complete power series solution of $y'' + x^2y' + 2xy = 0$ about the ordinary point $x = 0$. Your answer should include nice formulas for arbitrary coefficients in each of the fundamental solutions. You need to both set-up and solve the reccurrence relations as best you can.

PP 231 (Ritger & Rose 7-2 problem 7 part c) Find the first four nonzero terms in the power series solution about zero for the initial value problem $(x+2)y'' + 3y = 0$ with $y(0) = 0$ and $y'(0) = 1$.

PP 232 (Ritger & Rose 7-2 problem 7 part d) Find the first four nonzero terms in the power series solution about zero for the initial value problem $y'' + \sin(x)y' + (x-1)y = 0$ with $y(0) = 1$ and $y'(0) = 0$.

PP 233 Construct a differential equation with $y_1(x) = \frac{\sin(x)}{x}$ for $x \neq 0$ and $y_1(0) = 1$, $y_2(x) = x$ as its fundamental solution set. To accomplish this task do two tasks:

(a.) Argue from appropriate facts from the theory of determinants that $L[y] = \det \begin{bmatrix} y & y' & y'' \\ y_1 & y_1' & y_1'' \\ y_2 & y_2' & y_2'' \end{bmatrix}$ is a linear ODE with solutions y_1 and y_2.

(b.) calculate $L[y]$ explicitly as a linear ODE of the form $py'' + qy' + ry = 0$ where p, q, r are perhaps given as Taylor expansions about zero.

PP 234 (from page 103 of Boyce and DiPrima's 3rd Ed.) Consider $xy'' - (x+N)y' + Ny = 0$ for $N \in \mathbb{N}$

(a.) show $y_1 = e^x$ is a solution.

(b.) show that $y_2 = ce^x \int x^N e^{-x}\,dx$ is a second solution. (perhaps use the result of the previous problem, or the theorem from my notes or Ritger & Rose)

(c.) set $c = \frac{-1}{N!}$ and show by induction that $y_2(x) = T_n(x)$ the n-th order Taylor polynomial of e^x.

PP 235 (introduction to theory of adjoints, from page 95 of Boyce and DiPrima's 3rd Ed.) If $p(x)y'' + q(x)y' + r(x)y = 0$ can be expressed as $[p(x)y']' + [f(x)y]' = 0$ then it is said to be **exact**. Omit x-dependence in p, q, r, μ for brevity, if $py'' + qy' + ry = 0$ is not exact then it is possible to make it exact with multiplication by the appropriate integrating factor μ. **Show** that for μ to accomplish its stated task it must itself be the solution of the so-called **adjoint equation**

$$p\mu'' + (2p' - q)\mu' + (p'' - q' + r)\mu = 0.$$

where we have assumed p, q possess the stated derivatives. Find the adjoint equation for

a. [constant coefficient case] $ay'' + by' + cy = 0$

b. [Bessel Eqn. of order ν] $x^2y'' + xy' + (x^2 - \nu^2)y = 0$

c. [The Airy Eqn.] $y'' - xy = 0$

Chapter 8

partial differential equations

8.1 overview

In this chapter we study a particular solution techinque for partial differential equations. Our work here is certainly not comprehensive or self-contained. There are many simple modifications of the problems we consider which cannot be solved by the tools which are explicitly considered here. However, much of what is known beyond this chapter is a more or less simple twist of the methods considered here. Probably the most serious transgression is the complete lack of a complete treatment of Fourier series analysis. You can consult Chapters 13 and 14 of Ritger & Rose or Chapters 10 and 11 of Nagel, Saff and Snider's text. The Sturm-Lioville results are of particular interest and we have nothing to say about them here except that you ought to study them if you wish to go to the next step past this course. I'll begin with an overview of the techinque:

(i.) we are given a partial differential equation paired with a **boundary condition** (BC) and an **initial condition** (IC)

(ii.) we propose a solution which is formed by multiplying functions of just one variable.

(iii.) the proposed solution is found to depend on some characteristic value which (for reasons explained elsewhere) will depend on some integer n. Each solution u_n solves the given BCs.

(iv.) the initial condition is a function of at least one variable. To fit the initial condition we express it as a series of the u_n solutions. The Fourier techinque allows for elegant selection of the coefficients in the series.

(v.) the calculated coefficients are placed back in the formal solution and the solution depending on two or more variables is complete.

In the case of Laplace's equation, there are only BVs, but still the Fourier technique is needed to fit them together.

8.2 Fourier technique

There are further calculations given in PH127-133 which are linked as pdfs on our course website.

Let me record the basic formulas to calculate Fourier series here: first let me set the stage with some calculus II results you may have forgotten:

$$\int_{-L}^{L} \sin\left(\frac{m\pi x}{L}\right)\cos\left(\frac{n\pi x}{L}\right)\, dx = 0$$

$$\int_{-L}^{L} \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right)\, dx = L\delta_{m,n}$$

$$\int_{-L}^{L} \cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{n\pi x}{L}\right)\, dx = L\delta_{m,n} + L\delta_{m,0}\delta_{n,0}$$

The last formula indicates the case $m = n = 0$ is special for cosine, in that case the integral is just $\int_{-L}^{L} dx = 2L$. Generally $\delta_{m,n} = 1$ when $m = n$ and $\delta_{m,n} = 0$ when $m \neq n$, this is the **Kronecker delta**. Notice if we define:

$$\langle f(x), g(x)\rangle = \int_{-L}^{L} f(x)g(x)\, dx \qquad \& \qquad \|f(x)\| = \sqrt{\langle f(x), f(x)\rangle}.$$

Then $\langle\ ,\ \rangle$ behaves like a dot-product for functions and $\|\cdot\|$ gives a way for us to measure the *length* of the function $f(x)$. Notice the identities above show that for any m, n,

$$\left\langle \sin\left(\frac{m\pi x}{L}\right),\ \cos\left(\frac{n\pi x}{L}\right)\right\rangle = 0$$

and for $m \neq n$,

$$\left\langle \sin\left(\frac{m\pi x}{L}\right),\ \sin\left(\frac{n\pi x}{L}\right)\right\rangle = 0,$$

$$\left\langle \cos\left(\frac{m\pi x}{L}\right),\ \cos\left(\frac{n\pi x}{L}\right)\right\rangle = 0.$$

In other words, sine and cosine are orthogonal and sines with differing arguments are orthogonal and cosines with differing arguments are orthogonal. It turns out that for many functions we can use sines and cosines as a sort of basis to build the functions from a sum of sines or cosines. However, I say *basis* with some caution since we're not talking about finite linear combinations for most examples. Usually it takes an infinite series of sines or cosines or both in order to reproduce the given function. There are three main results we should know about for our future work in solving PDEs:

Notice the notation $f(x) \sim g(x)$ means that $f(x) = g(x)$ **almost everywhere**. This is given a precise meaning in **measure theory**. We can formally derive the results which follow from the integrals which I gave above.

Proposition 8.2.1. *Fourier Series on $[-L, L]$:*

Let f be piecewise continuous function on $[-L, L]$. Then

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right]$$

where the coefficients of the **trigonometric series** above are found via:

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx, \quad n = 0, 1, 2, \ldots$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx, \quad n = 1, 2, \ldots$$

.

There are also expansions which are special to $[0, L]$.

Proposition 8.2.2. *Cosine and Sine Series on $[0, L]$: (the* **half-range expansions***)*

Let f be piecewise continuous function on $[0, L]$. Then the **Fourier cosine series** of $f(x)$ on $[0, L]$ is

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right) \qquad \& \qquad a_n = \frac{2}{L} \int_{0}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx, \quad n = 0, 1, 2, \ldots.$$

Likewise, the **Fourier sine series** of $f(x)$ on $[0, L]$ is

$$f(x) \sim \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{L}\right) \qquad \& \qquad b_n = \frac{2}{L} \int_{0}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx, \quad n = 1, 2, \ldots$$

.

Example 8.2.3. *If $f(x) = x$ on $-L \leq x \leq L$ with $L = \pi$ then we can calculate:*

$$f(x) \sim \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx).$$

Likewise, if we calculate the sine series for $f(x)$ on $[0, \pi]$ is:

$$f(x) \sim \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx).$$

However, the cosine series for $f(x)$ on $[0, \pi]$

$$f(x) \sim \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} \cos((2k+1)x).$$

If we instead look at $f(x) = x^2$ then we would find the Fourier series and sine series would differ whereas the cosine series would mirror the full Fourier series. The sine series expansion gives an odd-extension of $f(x)$ given on $[0, L]$ whereas the cosine series expansion gives an even-extension of $f(x)$ given on $[0, L]$.

8.3 boundary value problems

In this section I treat the three main cases which arise in our study of PDEs. Here we consider families of ODEs, we'll see that only certain members of the family actually permit a solution of the given **boundary values**. In the interest of streamlining our work in the future sections we intend to refer to the work completed in this section. In each subsection that follows I present a calculation which justifies the proposition at the conclusion of the subsection.

8.3.1 zero endpoints

> **Problem:** determine which values of K allow solutions of $y'' + Ky = 0$ subject $y(0) = y(L) = 0$. We suppose x is the independent variable and we seek solutions for $0 \leq x \leq L$.

There are three cases to consider.

Case I: suppose $K = 0$. Then $y'' = 0$ has solution $y = Ax + B$ and thus $y(0) = B = 0$ and $y(L) = AL = 0$ hence $y = 0$ is the only solution in this case.

Case II: suppose $K < 0$. Then, there exists $\beta > 0$ for which $K = -\beta^2$. We face $y'' - \beta^2 y = 0$ thus solutions have the form $y = A\cosh(\beta x) + B\sinh(\beta x)$. Observe $y(0) = 0$ yields $0 = A$ hence $y(L) = 0$ yields $B\sinh(\beta L) = 0$. Hyperbolic sine is only zero at zero thus $\beta L = 0$. However, $\beta > 0$ hence there is no solution in this case.

Case III: suppose $K > 0$. Then, there exists $\beta > 0$ for which $K = \beta^2$. We face $y'' + \beta^2 y = 0$ thus solutions have the form $y = A\cos(\beta x) + B\sin(\beta x)$. Note, $y(0) = A$ hence $y(L) = B\sin(\beta L) = 0$. In considerable contrast to case II, the condition above permits infinitely many solutions:

$$\beta L = n\pi$$

for $n \in \mathbb{Z}$. However, as $\beta > 0$ we need only consider $n \in \mathbb{N}$ thus:

Proposition 8.3.1.

> The nontrivial solutions of $y'' + Ky = 0$ subject $y(0) = y(L) = 0$ have the form:
>
> $$y_n = B_n \sin\left(\frac{n\pi x}{L}\right) \qquad \text{for } n = 1, 2, 3, \ldots.$$
>
> Moreover, nontrivial solutions exist only if $K = \frac{n^2\pi^2}{L^2}$.

8.3.2 zero-derivative endpoints

> **Problem:** determine which values of K allow solutions of $y'' + Ky = 0$ subject the conditions $y'(0) = y'(L) = 0$. We suppose x is the independent variable and we seek solutions for $0 \leq x \leq L$.

There are three cases to consider.

Case I: suppose $K = 0$. Then $y'' = 0$ has solution $y = Ax + B$ and thus $y'(0) = A = 0$ and $y'(L) = A = 0$ hence $y = B$ is the only solution in this case. We do find one constant solution here which could be nontrivial.

Case II: suppose $K < 0$. Then, there exists $\beta > 0$ for which $K = -\beta^2$. We face $y'' - \beta^2 y = 0$ thus solutions have the form $y = A\cosh(\beta x) + B\sinh(\beta x)$. Observe $y'(0) = 0$ yields $0 = \beta B$ hence $y'(L) = 0$ yields $\beta A\sinh(\beta L) = 0$. Hyperbolic sine is only zero at zero thus $\beta L = 0$. However, $\beta > 0$ hence there is no solution in this case.

Case III: suppose $K > 0$. Then, there exists $\beta > 0$ for which $K = \beta^2$. We face $y'' + \beta^2 y = 0$ thus solutions have the form $y = A\cos(\beta x) + B\sin(\beta x)$. Note, $y'(0) = 0$ yields $\beta B = 0$ hence $y'(L) = -\beta A\sin(\beta L) = 0$. This condition permits infinitely many solutions:

$$\beta L = n\pi$$

for $n \in \mathbb{Z}$. However, as $\beta > 0$ we need only consider $n \in \mathbb{N}$ thus (including the constant solution as $n = 0$ in what follows below) we find:

Proposition 8.3.2.

The nontrivial solutions of $y'' + Ky = 0$ subject $y'(0) = y'(L) = 0$ have the form:

$$y_n = A_n\cos\left(\frac{n\pi x}{L}\right) \qquad \text{for } n = 0, 1, 2, 3, \dots.$$

Moreover, nontrivial solutions exist only if $K = \frac{n^2\pi^2}{L^2}$ for some $n \in \mathbb{N} \cup \{0\}$.

8.3.3 mixed endpoints

> **Problem:** determine which values of K allow solutions of $y''+Ky = 0$ subject $y(0) = 0$ and $y'(L) = 0$. We suppose x is the independent variable and we seek solutions for $0 \leq x \leq L$.

There are three cases to consider.

Case I: suppose $K = 0$. Then $y'' = 0$ has solution $y = Ax + B$ and thus $y(0) = B = 0$ and $y'(L) = A = 0$ hence $y = 0$ is the only solution in this case.

Case II: suppose $K < 0$. Then, there exists $\beta > 0$ for which $K = -\beta^2$. We face $y'' - \beta^2 y = 0$ thus solutions have the form $y = A\cosh(\beta x) + B\sinh(\beta x)$. Observe $y(0) = 0$ yields $A = 0$ hence $y'(L) = \beta B\cosh(\beta L) = 0$ yields $B = 0$. There is no solution in this case.

Case III: suppose $K > 0$. Then, there exists $\beta > 0$ for which $K = \beta^2$. We face $y'' + \beta^2 y = 0$ thus solutions have the form $y = A\cos(\beta x) + B\sin(\beta x)$. Note, $y(0) = 0$ yields $A = 0$ hence $y'(L) = \beta B\cos(\beta L) = 0$. This condition permits infinitely many solutions:

$$\beta L = (2n-1)\frac{\pi}{2}$$

for $n \in \mathbb{Z}$. However, as $\beta > 0$ we need only consider $n \in \mathbb{N}$ thus:

Proposition 8.3.3.

The nontrivial solutions of $y'' + Ky = 0$ subject $y(0) = y'(L) = 0$ have the form:

$$y_n = B_n \sin\left(\frac{\pi(2n-1)x}{2L}\right) \qquad \text{for } n = 1, 2, 3, \dots.$$

Moreover, nontrivial solutions exist only if $K = \frac{\pi^2(2n-1)^2}{4L^2}$ for some $n \in \mathbb{N}$.

Of course, there is one more mixed case to consider:

> **Problem:** determine which values of K allow solutions of $y''+Ky = 0$ subject $y'(0) = 0$ and $y(L) = 0$. We suppose x is the independent variable and we seek solutions for $0 \leq x \leq L$.

There are again three cases to consider. I omit the details, but, once again $K = 0$ and $K < 0$ add no nontrivial solutions. On the other hand, when $K = \beta^2 > 0$ we have solutions of the form:

$$y = A\cos(\beta x) + B\sin(\beta x)$$

Note, $y'(0) = \beta B = 0$ thus $y(L) = 0$ yields $A\cos(\beta L) = 0$. This condition permits infinitely many solutions:

$$\beta L = (2n-1)\frac{\pi}{2}$$

for $n \in \mathbb{Z}$. However, as $\beta > 0$ we need only consider $n \in \mathbb{N}$ thus:

Proposition 8.3.4.

The nontrivial solutions of $y'' + Ky = 0$ subject $y'(0) = y(L) = 0$ have the form:

$$y_n = A_n \cos\left(\frac{\pi(2n-1)x}{2L}\right) \qquad \text{for } n = 1, 2, 3, \dots.$$

Moreover, nontrivial solutions exist only if $K = \frac{\pi^2(2n-1)^2}{4L^2}$ for some $n \in \mathbb{N}$.

Remark 8.3.5.

By now you might wonder why we even bother with hyperbolic sine and cosine in these boundary value problems. It seems only sine, cosine and the constant solution appear nontrivially. Indeed that is the take-away message of this section. For trivial or unchanging endpoints we find either sine, cosine or constant solutions. That said, logic requires us to investigate all cases at least to begin our study. Also, you will see hyperbolic sine and/or cosine appear in the solution to Laplace's equation on some rectangular domain.

8.4 heat equations

In this section we discuss how one may solve the one-dimensional heat equation by the technique of separation of variables paired with the Fourier technique.

> **Problem:** Let $u = u(t,x)$ denote the temperature u at position x and time t along an object from $x = 0$ to $x = L$. The time-evolution of temperature in such an object is modelled by
> $$u_{xx} = \alpha u_t$$
> where α is a constant determined by the particular physical characteristics of the object. If we are given an initital temperature distribution $f(x)$ for $0 \leq x \leq L$ and the system is subject to boundary conditions at $x = 0, L$ of the type studied in Propositions 8.3.1, 8.3.2, 8.3.3, and 8.3.4 then **find** $u(x,t)$ **for** $t > 0$ **and** $0 \leq x \leq L$.

Let us study the general problem

$$\boxed{u_{xx} = \alpha u_t \qquad \star}$$

and see what we can conjecture independent of a particular choice of BV and IC[1]. Let us suppose the solution of $\star$ can be written as product of functions of x and t:

$$u(x,t) = X(x)T(t)$$

Observe, is we omit arguments and use $X' = \frac{dX}{dx}$ and $T' = \frac{dT}{dt}$ then:

$$u_x = X'T, \qquad u_{xx} = X''T, \qquad u_t = XT'$$

Subsitute the above into $\star$ to obtain:

$$X''T = \alpha XT'$$

Therefore, dividing by XT yields:

$$\frac{X''}{X} = \alpha \frac{T'}{T}$$

This is an interesting equation as the left and right hand sides are respectively functions of only x or only t. There is only one way this can happen; both sides must be constant. In other words, there must exist some constant $-K$ for which

$$\frac{X''}{X} = -K = \alpha \frac{T'}{T}$$

Hence, we find **two** families of ODEs which are linked through this characteristic value K. We would like to solve the following simultaneously:

$$X'' + KX = 0 \qquad \& \qquad T' = (-K/\alpha)T$$

The time function is easily solved by separation of variables to give $T(t) = T_o \exp(-Kt/\alpha)$ where $T_o \neq 0$ and customarily[2] we may set $T_o = 1$. At this point we cannot go further in detail without

[1]in case you missed it, BV is boundary value and here this indicates $u = 0$ or $u_x = 0$ at $x = 0, L$ for all $t > 0$ and IC is initial condition which we denote $f(x)$ where we suppose $u(x,0) = f(x)$ for $0 \leq x \leq L$

[2]why this is reasonable is clear later in the calculation, we introduce constants for the formal solution of X thus any multiplicative factor in T can be absorbed into those constants

specifying the details of the boundary values. I must break into cases here.

Case I. (ends at zero temperature) suppose $u(0,t) = u(L,t) = 0$ for $t > 0$. This indicates $X(0)T(t) = X(L)T(t) = 0$ hence, as $T(t) \neq 0$ for $t > 0$ we find $X(0) = X(L) = 0$ hence, by Proposition 8.3.1:

$$X_n(x) = B_n \sin\left(\frac{n\pi x}{L}\right)$$

for $n = 1, 2, 3, \dots$. In this case $K = \frac{n^2\pi^2}{L^2}$ thus we find $T_n(t) = \exp\left(\frac{-n^2\pi^2 t}{\alpha L^2}\right)$. I include the subscript n here to reflect the fact that there is not just one T solution. In fact, for each n the corresponding K provides a T_n solution. It is wise to reflect the n-dependence on the product solution; let $u_n(x,t) = X_n(x)T_n(t)$. Thus,

$$u_n(x,t) = B_n \sin\left(\frac{n\pi x}{L}\right) \exp\left(\frac{-n^2\pi^2 t}{\alpha L^2}\right)$$

Each choice of n gives us a product solution which solves $\star$ subject the BVs $u(0,t) = u(L,t) = 0$ for $t > 0$.

Case II. (ends perfectly insulated) suppose $u_x(0,t) = u_x(L,t) = 0$ for $t > 0$. This indicates $X'(0)T(t) = X'(L)T(t) = 0$ hence, as $T(t) \neq 0$ for $t > 0$ we find $X'(0) = X'(L) = 0$. Proposition 8.3.2 provides

$$X_n(x) = A_n \cos\left(\frac{n\pi x}{L}\right)$$

where $K = \frac{n^2\pi^2}{L^2}$ thus $T_n(t) = \exp\left(\frac{-n^2\pi^2 t}{\alpha L^2}\right)$. The product solutions $u_n(x,t) = X_n(x)T_n(t)$ have the form

$$u_n(x,t) = A_n \cos\left(\frac{n\pi x}{L}\right) \exp\left(\frac{-n^2\pi^2 t}{\alpha L^2}\right).$$

Case III.(left end at zero temperature, right end insulated) suppose $u(0,t) = u_x(L,t) = 0$ for $t > 0$. This indicates $X(0)T(t) = X'(L)T(t) = 0$ hence, as $T(t) \neq 0$ for $t > 0$ we find $X(0) = X'(L) = 0$. Proposition 8.3.3 shows us

$$X_n(x) = B_n \sin\left(\frac{\pi(2n-1)x}{2L}\right) \qquad \text{for } n = 1, 2, 3, \dots.$$

Where $K = \frac{\pi^2(2n-1)^2}{4L^2}$. The dependence on n is once more reflected in K and thus $T(t) = \exp\left(\frac{-\pi^2(2n-1)^2 t}{4\alpha L^2}\right)$. The product solutions $u_n(x,t) = X_n(x)T_n(t)$ have the form

$$u_n(x,t) = B_n \sin\left(\frac{\pi(2n-1)x}{2L}\right) \exp\left(\frac{-\pi^2(2n-1)^2 t}{4\alpha L^2}\right).$$

Case IV.(left end insulated, right end at zero temperature) suppose $u_x(0,t) = u(L,t) = 0$ for $t > 0$. I invite the reader to determine the product solution $u_n(x,t)$ in the same way as the previous three cases. Of course, in this case you will need to use Proposition 8.3.4.

Let us return once more to the abstract case. Suppose we choose one of the four boundary condition types as laid forth in Propositions 8.3.1, 8.3.2, 8.3.3, and 8.3.4. Let $u_n(x,t)$ be the product solution as discussed above in Cases I-IV for[3] $n = 0, 1, 2, \dots$. Then the general formal solution to $u_x x = \alpha u_t$ subject the given BVs is formed by:

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t).$$

I say **formal** as I have no intention of properly analyzing the convergence of the series above. That said, it is natural to form the solution to the linear PDE $u_{xx} - \alpha u_t = 0$ by superposition all the solutions we found. You might wonder, have we really found **all** the solutions? It turns out the answer is yes. Once again, the proof is beyond this course[4]. Everything we have done thus far required no initial condition be given. We shall see that the multitude of constants implicit within the formal solution above gives us the flexibility to match any initial condition which is compatible with the given BV problem. The Fourier technique provides the necessary formulas to calculate the coefficients in the formal solution and thus we find a unique solution to a given BV problem paired with an initial condition. Of course, without the initial condition there are **infinitely many** product solutions which are allowed by the given BV problem. Rather than attempt some over-arching Case I-IV description of the Fourier technique I will focus on just one choice and show how it works out in that case.

Example 8.4.1. *We worked through $u_t = 3u_{xx}$ in Lecture 4-17-2014. See PH 135-138 for details. (the solution may have an error there, it may be missing a factor of $1/\pi$ which causes the answer found to be incorrect, I link it here since perhaps the general work shown is still helpful)*

Example 8.4.2. Problem: *Suppose $u_t = 5u_{xx}$ for $0 \leq x \leq \pi$ and $t > 0$. Furthermore, we are given boundary conditions $u(0,t) = u(\pi,t) = 0$ for $t > 0$ and an initial condition $u(x,0) = 1 - \cos(2x)$ for $0 \leq x \leq \pi$*

Solution: *we are in Case I hence the n-th eigensolution has the form:*

$$u_n(x,t) = B_n \sin\left(\frac{n\pi x}{L}\right) \exp\left(\frac{-n^2\pi^2 t}{\alpha L^2}\right)$$

But, identify $L = \pi$ and $\alpha = 1/5$ in the terminology of this section. Therefore, we find total formal solution:

$$u(x,t) = \sum_{n=1}^{\infty} B_n \sin(nx) \exp\left(-5n^2 t\right)$$

Next we fit the initial condition $u(x,0) = 1 - \cos(2x)$. Notice the exponential function reduces to 1 as we evaluate at $t = 0$. It follows that We need

$$1 - \cos(2x) = \sum_{n=1}^{\infty} B_n \sin(nx)$$

Unfortunately, the lhs is the wrong kind of Fourier series, so, we must calculate the sine-series for the function $f(x) = 1 - \cos(2x)$. Recall the Fourier sine series for $[0, \pi]$

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx) \qquad \& \qquad b_n = \frac{2}{\pi}\int_0^{\pi} f(x)\sin(nx)\,dx.$$

[3] we throw in $X_0 = 0$ for the cases which started at $n = 1$ to keep this discussion unfragmented

[4] I can recommend a few texts beyond these notes and your required text which face the needed analysis to give meaningful critques of these issues

Thus, we calculate: (see PH134 for integration details, only n odd is nontrivial, $b_{2k} = 0$ for $k = 1, 2, \dots$)

$$b_{2k-1} = \frac{2}{\pi}\int_0^{\pi} \big(1-\cos(2x)\big)\sin((2k-1)x)\,dx$$
$$= \frac{2}{\pi}\left[\frac{2}{2k-1} - \frac{1}{2k+1} + \frac{1}{3-2k}\right].$$

At this point, we have two[5] Fourier series:

$$\underbrace{\sum_{n=1}^{\infty} b_n \sin(nx)}_{from\ Fourier\ analysis} = \underbrace{\sum_{n=1}^{\infty} B_n \sin(nx)}_{from\ BV\ analysis}$$

Now, equation coefficients to see $b_n = B_n$ for all n. Thus, the solution is simply:

$$u(x,t) = \sum_{k=1}^{\infty} \frac{2}{\pi}\left[\frac{2}{2k-1} - \frac{1}{2k+1} + \frac{1}{3-2k}\right] \sin((2k-1)x)\exp\left(-5(2k-1)^2 t\right).$$

Example 8.4.3. Problem: *Suppose $7u_t = u_{xx}$ for $0 \le x \le 1/2$ and $t > 0$. Furthermore, we are given boundary conditions $u(0,t) = u_x(1/2,t) = 0$ for $t > 0$ and an initial condition $u(x,0) = \sin(\pi x) + 42\sin(17\pi x)$ for $0 \le x \le 1/2$*

Solution: *we are in Case III hence the n-th eigensolution has the form:*

$$u_n(x,t) = B_n \sin\left(\frac{\pi(2n-1)x}{2L}\right)\exp\left(\frac{-\pi^2(2n-1)^2 t}{4\alpha L^2}\right).$$

Observe $L = 1/2$ and $\alpha = 7$ thus the formal solution to the BV problem is

$$u(x,t) = \sum_{n=1}^{\infty} B_n \sin(\pi(2n-1)x)\exp\left(\frac{-\pi^2(2n-1)^2 t}{7}\right).$$

Now we wish to choose $B_1, B_2, \dots$ to fit the given intial condition. This time the initial data is already presented as a finite Fourier series so life is easy: just compare the expressions below and equate coefficients of sine functions,

$$u(x,0) = \sin(\pi x) + 42\sin(17\pi x) = \sum_{n=1}^{\infty} B_n \sin(\pi(2n-1)x).$$

Evidentally $B_1 = 1$ whereas $B_9 = 42$ and $n \neq 1, 9$ yields $B_n = 0$. Therefore, we obtain the solution:

$$u(x,t) = \sin(\pi x)\exp\left(\frac{-\pi^2 t}{7}\right) + 42\sin(17\pi x)\exp\left(\frac{-289\pi^2 t}{7}\right).$$

[5] equivalently, we could anticipate this step and simply note the Fourier integrals necessarily calculate B_n. I include this here to help you see the separation between the BV part of the solution process and the Fourier technique. Calculationally, there is a shorter path.

8.5 wave equations

In this section we discuss how one may solve the one-dimensional wave equation by the technique of separation of variables paired with the Fourier technique.

> **Problem:** Let $y = y(x,t)$ denote the vertical position of a string at position (x, y) and time t for a length which is stretched from $x = 0$ to $x = L$. The motion of the string is modelled by:
> $$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}$$
> where $v > 0$ is a constant which describes the speed of the wave on the string. If we are given an initial shape for the string $y(x, 0) = f(x)$ for $0 \leq x \leq L$ and an initial velocity of the string $y_t(x, 0) = g(x)$ where the system is subject to boundary conditions $y(0,t) = y(L,t) = 0$ (this type studied in Proposition 8.3.1) then **find** $y(x,t)$ **for** $t > 0$ **and** $0 \leq x \leq L$.

I should mention, while I have motivated this problem partly as the problem of a vibrating string the math we cover below equally well applies to electric or magnetic waves in the vacuum and a host of other waves which are transversely propagated. I also should confess, this section is a bit more narrow in focus than the previous section. I consider only the fixed endpoint case.

Let us begin by stating the problem: solve

$$\boxed{v^2 y_{xx} = y_{tt}. \qquad \star}$$

subject the boundary conditions $y(0,t) = 0$ and $y(L,t) = 0$ and subject the initial conditions $y(x, 0) = f(x)$ and $y_t(x, 0) = g(x)$ for $0 \leq x \leq L$. In contrast to the heat equation, we have two initial conditions to match to our solution. We begin in the same manner as the last section. We propose the solution separates into product solutions $y(x,t) = X(x)T(t)$. We omit arguments and use $X' = \frac{dX}{dx}$ and $T' = \frac{dT}{dt}$ then:

$$y_{xx} = X''T, \qquad y_{tt} = XT''$$

Subsitute the above into $\star$ to obtain:

$$v^2 X''T = XT''$$

Therefore, dividing by XT yields:

$$\frac{X''}{X} = \frac{1}{v^2}\frac{T''}{T}$$

This is an interesting equation as the left and right hand sides are respectively functions of only x or only t. There is only one way this can happen; both sides must be constant. Once more, I'll denote this constant $-K$ hence

$$\frac{X''}{X} = \frac{1}{v^2}\frac{T''}{T} = -K \qquad \Rightarrow \qquad X'' + KX = 0, \;\; \& \;\; T'' + Kv^2T = 0.$$

Given $y(x, 0) = 0$ and $y(L, 0) = 0$ we once more conclude from the supposed nontriviality of $T(t)$ that $X(0) = 0$ and $X(L) = 0$ hence we may apply Proposition 8.3.1 to find the family of eigensolutions for X:

$$X_n(x) = \sin\left(\frac{n\pi x}{L}\right)$$

where we find that $K = \frac{n^2\pi^2}{L^2}$ for $n \in \mathbb{N}$. The time equation also involves the same K. Thus, we solve:

$$T'' + \frac{v^2 n^2 \pi^2}{L^2} T = 0$$

Let $\beta = \frac{\pi n v}{L}$ then it is clear $T'' + \beta^2 T = 0$ has solutions (for each n)

$$T_n(t) = A_n \cos(\beta t) + B_n \sin(\beta t).$$

In total, we find the n-eigensolution to the BV problem for the wave-equation with fixed ends is:

$$y_n(x,t) = A_n \sin\left(\frac{n\pi x}{L}\right)\cos(\beta t) + B_n \sin\left(\frac{n\pi x}{L}\right)\sin(\beta t)$$

Thus, the general formal solution of the problem is given by:

$$\boxed{y(x,t) = \sum_{n=1}^{\infty}\left[A_n \sin\left(\frac{n\pi x}{L}\right)\cos(\beta t) + B_n \sin\left(\frac{n\pi x}{L}\right)\sin(\beta t)\right] \quad \star\star}$$

The initial conditions will force us to select particular values for the coefficients A_n and B_n. In particular, notice how setting $t = 0$ makes the B_n terms vanish. Thus, from $\star\star$ we see $f(x) = y(x,0)$ yields:

$$f(x) = y(x,0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right)$$

therefore, we will need to find a Fourier sine series for $f(x)$ on $0 \leq x \leq L$ to fit the coefficients A_n from the particular shape of $f(x)$. On the other hand, when we calculate $\frac{\partial y}{\partial t}(x,0)$ we find the A_n term vanishes due to a sine evaluated at zero and the differentiation and evaluation leaves: (applying $g(x) = u_t(x,0)$ to $\star\star$)

$$g(x) = \sum_{n=1}^{\infty} \beta B_n \sin\left(\frac{n\pi x}{L}\right) = \sum_{n=1}^{\infty} \frac{\pi n v}{L} B_n \sin\left(\frac{n\pi x}{L}\right).$$

Once given a particular $g(x)$ we again see that a Fourier sine series for $g(x)$ on $0 \leq x \leq L$ will solve the problem. However, note B_n are not the same as the coefficients in the Fourier series. Rather, if the Fourier expansion for $g(x)$ is found to be $g(x) = \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right)$ then we need to choose B_n to solve:

$$C_n = \frac{\pi n v}{L} B_n.$$

That choice will fit the given initial velocity for the wave and paired with the correct Fourier analysis for A_n will produce a solution to the given wave equation.

Example 8.5.1. *See PH 145-Ph148 for a complete solution to a wave equation from front to back. Notice, if you use what we have derived thus far in this section then your work is streamlined a bit in comparison.*

Example 8.5.2. Problem: *Suppose $y_{tt} = 4y_{xx}$ for $0 \leq x \leq \pi$ and $t > 0$. Furthermore, we are given boundary conditions $y(0,t) = y(\pi,t) = 0$ for $t > 0$ and an initial conditions $y(x,0) = x^2(\pi - x)$ and $y_t(x,0) = 0$ for $0 < x < \pi$.*

Solution: *Note $v = 2$ and $L = \pi$ and $\beta = 2n$ for the given problem hence the general solution $\star\star$ takes the form:*

$$y(x,t) = \sum_{n=1}^{\infty} \left[A_n \sin(nx)\cos(2nt) + B_n \sin(nx)\sin(2nt)\right]$$

Observe $y_t(x,0) = 0$ yields $B_n = 0$ for all $n \in \mathbb{N}$. On the other hand, the coefficients A_n may be determined from the Fourier expansion of $f(x) = x^2(\pi - x)$ on $0 \leq x \leq \pi$ as a sine series. If $f(x) = \sum_{n=1}^{\infty} a_n \sin(nx)$ then we can calculate by the sine-series Fourier coefficient formula

$$a_n = \frac{2}{\pi}\int_0^{\pi} x^2(\pi - x)\sin(nx)\, dx = \frac{4}{n^2}\left[2(-1)^{n+1} - 1\right].$$

To fit $y(x,0) = x^2(\pi - x)$ we simply set $A_n = a_n$ hence the solution is simply:

$$y(x,t) = \sum_{n=1}^{\infty} \frac{4}{n^2}\left[2(-1)^{n+1} - 1\right]\sin(nx)\cos(2nt).$$

Example 8.5.3. Problem: *Suppose $y_{tt} = 4y_{xx}$ for $0 \leq x \leq \pi$ and $t > 0$. Furthermore, we are given boundary conditions $y(0,t) = y(\pi,t) = 0$ for $t > 0$ and an initial conditions $y(x,0) = \sin(x) + 13\sin(4x)$ and $y_t(x,0) = 7\sin(5x)$ for $0 < x < \pi$.*

Solution: *Note $v = 2$ and $L = \pi$ and $\beta = 2n$ for the given problem hence the general solution $\star\star$ takes the form:*

$$y(x,t) = \sum_{n=1}^{\infty} \left[A_n \sin(nx)\cos(2nt) + B_n \sin(nx)\sin(2nt)\right]$$

Plug in the initial condition $y(x,0) = \sin(x) + 13\sin(4x)$ to obtain:

$$\sin(x) + 13\sin(4x) = \sum_{n=1}^{\infty} A_n \sin(nx)$$

we find $A_1 = 1$, $A_4 = 13$ and otherwise $A_n = 0$. Differentiate and evaluate $y(x,t)$ at $t = 0$ to apply the initial velocity condition to the general solution:

$$7\sin(5x) = \sum_{n=1}^{\infty} 2nB_n \sin(nx).$$

we find $7 = 10B_5$ thus $B_5 = 7/10$ and $B_n = 0$ for $n \neq 5$. In total, the given initial conditions leave just three nontrivial terms in the infinite sum which formed the general solution. We find:

$$y(x,t) = \sin(x)\cos(2t) + 13\sin(4x)\cos(8t) + \frac{7}{10}\sin(5x)\sin(10t).$$

It is interesting to animate the solution on $0 < x < \pi$ to see how the string waves. Finally, I should mention there is much more to learn about the wave equation. For waves on an infinite string there is even a simple solution with no need for Fourier analysis. See PH148-PH149 for an example of D'Almbert's solution.

8.6 Laplace's equation

Laplace's equation arises in the study of fluid flow as well as electrostatics. In the context of electrostatics ϕ is the voltage. Indeed, it is one of the fundamental problems of mathematical physics to solve Laplace's equation in various contexts.

Problem: Let R be a rectangle in the xy-plane. We seek to solve

$$\phi_{xx} + \phi_{yy} = 0$$

given various conditions for ϕ and $\nabla\phi$ on ∂R.

We just study an example here. I make no claim of generality. That said, what we do here is largely analogous to our work on the heat and wave equation. The main idea is separation of variables.

Example 8.6.1. Problem: *Let $R = [0, \pi] \times [0, 1]$ be a rectangle in the xy-plane. We seek to solve*

$$\phi_{xx} + \phi_{yy} = 0$$

subject the boundary conditions: $\phi(x, 1) = 0$ and $\phi(x, 0) = f(x)$ for $x \in [0, \pi]$ and $\phi(0, y) = \phi(\pi, y) = 0$ for $y \in [0, 1]$. I leave $f(x)$ unspecified for the moment.

Solution: *we propose $\phi(x, y) = X(x)Y(y)$ and note $\phi_{xx} = X''Y$ whereas $\phi_{yy} = XY''$. Thus,*

$$X''Y + XY'' = 0 \quad \Rightarrow \quad \frac{X''}{X} = -\frac{Y''}{Y} = -K \quad \Rightarrow \quad X'' + KX = 0, \; Y'' - KY = 0.$$

Again, the only option above is that both quotients be constant as the l.h.s. is just a function of x whereas the r.h.s. is just a function of y. Notice that $\phi(0, y) = X(0)Y(y) = 0$ and $\phi(\pi, y) = X(\pi)Y(y) = 0$. Thus $X(0) = X(\pi) = 0$ hence we apply Proposition 8.3.1 to find the family of eigensolutions for X:

$$X_n(x) = \sin(nx)$$

where we find that $K = n^2$ for $n \in \mathbb{N}$. Thus $Y'' - n^2 Y = 0$ and we find[6] eigensolutions for Y

$$Y_n(y) = A_n \sinh(ny + C_n)$$

We also have $\phi(x, 1) = 0$ hence $X(x)Y(1) = 0$

$$Y_n(1) = A_n \sinh(n + C_n) = 0$$

hence $n + C_n = 0$ and we find $Y_n(y) = A_n \sinh(n(y - 1))$. We form the general solution which solves all but the last boundary condition in terms of a formal sum over the product solutions $X_n(x)Y_n(y)$.

$$\phi(x, y) = \sum_{n=1}^{\infty} A_n \sin(nx) \sinh(n(y - 1)).$$

The remaining boundary condition involves $f(x)$. In particular, $\phi(x, 0) = f(x)$ yields

$$f(x) = \phi(x, 0) = \sum_{n=1}^{\infty} A_n \sin(nx) \sinh(-n)$$

Thus, to solve such a problem we need only find the Fourier expansion for $f(x) = \sum_{n=1}^{\infty} c_n \sin(nx)$ and observe that $c_n = -\sinh(n)A_n$ so we can easily assemble the solution.

[6] here I follow Ritger and Rose page 466 where a similar problem is solved, alternatively you could work this out with $Y_n(y) = A_n \cosh(ny) + B_n \sinh(ny)$

Example 8.6.2. *See PH 151-155 for more explicit calculation of Laplace Equation solutions in Cartesian coordinates.*

In PH157-161 I explore how to solve Laplace's Equation in polar coordinates. Indeed, there are many more techniques which are known for this problem. In particular, in complex analysis one learns how conformal mapping allows elegant solutions of problems far more complicated than those we consider here.

Chapter 9

introduction to variational calculus

9.1 history

The problem of variational calculus is almost as old as modern calculus. Variational calculus seeks to answer questions such as:

Remark 9.1.1.

1. what is the shortest path between two points on a surface ?
2. what is the path of least time for a mass sliding without friction down some path between two given points ?
3. what is the path which minimizes the energy for some physical system ?
4. given two points on the x-axis and a particular area what curve has the longest perimeter and bounds that area between those points and the x-axis?

You'll notice these all involve a variable which is not a real variable or even a vector-valued-variable. Instead, the answers to the questions posed above will be **paths** or **curves** depending on how you wish to frame the problem. In variational calculus the variable is a function and we wish to find extreme values for a **functional**. In short, a functional is an abstract function of functions. A functional takes as an input a function and gives as an output a number. The space from which these functions are taken varies from problem to problem. Often we put additional **contraints** or **conditions** on the **space of admissable solutions**. To read about the full generality of the problem you should look in a text such as Hans Sagan's. Our treatment is introductory in this chapter, my aim is to show you why it is plausible and then to show you how we use variational calculus.

We will see that the problem of finding an extreme value for a functional is equivalent to solving the Euler-Lagrange equations or Euler equations for the functional. Euler predates Lagrange in his discovery of the equations bearing their names. Eulers's initial attack of the problem was to chop the hypothetical solution curve up into a polygonal path. The unknowns in that approach were the coordinates of the vertices in the polygonal path. Then through some ingenious calculations he arrived at the Euler-Lagrange equations. Apparently there were logical flaws in Euler's original treatment. Lagrange later derived the same equations using the viewpoint that the variable was a function and the **variation** was one of shifting by an arbitrary function. I'm no expert of the full history, I just give you a rough sketch of what I've gathered from reading a few variational calculus texts.

Physics played a large role in the development of variational calculus. Lagrange was a physicist as well as a mathematician. At the present time, every physicist takes course(s) in *Lagrangian Mechanics*. Moreover, the use of variational calculus is fundamental since Hamilton's principle says that all physics can be derived from the principle of least action. In short this means that nature is lazy. The solutions realized in the physical world are those which minimize the action. The action

$$S[y] = \int L(y, y', t)\, dt$$

is constructed from the Lagrangian $L = T - U$ where T is the kinetic energy and U is the potential energy. In the case of classical mechanics the Euler Lagrange equations are precisely Newton's equations. The Hamiltonian $H = T + U$ is similar to the Lagrangian except that the fundamental variables are taken to be momentum and position in contrast to velocity and position in Lagrangian mechanics.

Hamiltonians and Lagrangians are used to set-up new physical theories. Euler-Lagrange equations are said to give the so-called *classical limit* of modern field theories. The concept of a force is not so useful to quantum theories, instead the concept of energy plays the central role. Moreover, the problem of quantizing and then renormalizing field theory brings in very sophisiticated mathematics. In fact, the math of modern physics is not understood. In this chapter I'll just show you a few famous classical mechanics problems which are beatifully solved by Lagrange's approach. We'll also see how expressing the Lagrangian in non-Cartesian coordinates can give us an easy way to derive forces that arise from geometric contraints.

I am following the typical physics approach to variational calculus. The treatment given here is close to that of Arfken and Weber's Mathematical Physics text, however I suspect you can find these calculations in dozens of classical mechanics texts. More or less our approach is that of Lagrange.

I begin with a derivation of the Euler-Lagrange equation for one-dimensional problems and I present solutions to several standard problems. Then we turn to finding the Euler-Lagrange equations for problems with many dependent variables. I will not cover variational calculus derviations of partial differential equations in this course. If you're curious, I discuss how to derive Maxwell's Equations via such calculus here http://www.supermath.info/Lecture10_ClassicalFieldTheory.pdf

9.2 the variational problem

Our goal in what follows here is to maximize or minimize a particular function of functions. Suppose $\mathcal{F}_o$ is a set of functions with some particular property. For now, we may could assume that all the functions in $\mathcal{F}_o$ have graphs that include (x_1, y_1) and (x_2, y_2). Consider a functional $J : \mathcal{F}_o \to \mathcal{F}_o$ which is defined by an integral of some function f which we call the **Lagrangian**,

$$J[y] = \int_{x_1}^{x_2} f(y, y', x)\, dx.$$

We suppose that f is given but y is a variable. Consider that if we are given a function $y^* \in \mathcal{F}_o$ and another function η such that $\eta(x_1) = \eta(x_2) = 0$ then we can reach a whole family of functions indexed by a real variable α as follows (relabel $y^*(x)$ by $y(x, 0)$ so it matches the rest of the family of functions):

$$y(x, \alpha) = y(x, 0) + \alpha\eta(x)$$

Note that $x \mapsto y(x, \alpha)$ gives a function in $\mathcal{F}_o$. We define the **variation** of y to be

$$\boxed{\delta y = \alpha\eta(x)}$$

This means $y(x, \alpha) = y(x, 0) + \delta y$. We may write J as a function of α given the variation we just described:

$$J(\alpha) = \int_{x_1}^{x_2} f(y(x, \alpha), y(x, \alpha)', x)\, dx.$$

It is intuitively obvious that if the function $y^*(x) = y(x, 0)$ is an extremum of the functional then we ought to expect

$$\left[\frac{\partial J(\alpha)}{\partial \alpha}\right]_{\alpha=0} = 0$$

Notice that we can calculate the derivative above using multivariate calculus. Remember that $y(x, \alpha) = y(x, 0) + \alpha\eta(x)$ hence $y(x, \alpha)' = y(x, 0)' + \alpha\eta(x)'$ thus $\frac{\partial y}{\partial \alpha} = \eta$ and $\frac{\partial y'}{\partial \alpha} = \eta' = \frac{d\eta}{dx}$. Consider that:

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[\int_{x_1}^{x_2} f(y(x, \alpha), y(x, \alpha)', x)\, dx\right] \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha} + \frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} \right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\frac{d\eta}{dx} \right) dx \qquad (9.1)
\end{aligned}$$

Observe that

$$\frac{d}{dx}\left[\frac{\partial f}{\partial y'}\eta\right] = \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta + \frac{\partial f}{\partial y'}\frac{d\eta}{dx}$$

Hence continuing Equation 9.1 in view of the product rule above,

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta + \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\eta\right] - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta \right) dx \\
&= \frac{\partial f}{\partial y'}\eta\Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta \right) dx \qquad (9.2) \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] \right)\eta\, dx
\end{aligned}$$

Note we used the conditions $\eta(x_1) = \eta(x_2)$ to see that $\frac{\partial f}{\partial y'}\eta\big|_{x_1}^{x_2} = \frac{\partial f}{\partial y'}\eta(x_2) - \frac{\partial f}{\partial y'}\eta(x_1) = 0$. Our goal is to find the extreme values for the functional J. Let me take a few sentences to again restate our set-up. Generally, we take a function y then J maps to a new function $J[y]$. The family of functions indexed by α gives a whole collection of functions in $\mathcal{F}_o$ which are near y^* according to the formula,

$$y(x, \alpha) = y^*(x) + \alpha\eta(x)$$

Let's call this set of functions W_η. If we took another function like η, say ζ such that $\zeta(x_1) = \zeta(x_2) = 0$ then we could look at another family of functions:

$$y(x,\alpha) = y^*(x) + \alpha\zeta(x)$$

and we could denote the set of all such functions generated from ζ to be W_ζ. The total variation of y based at y^* should include all possible families of functions in $\mathcal{F}_o$. You could think of W_η and W_ζ be two different subspaces in $\mathcal{F}_o$. If $\eta \neq \zeta$ then these subspaces of $\mathcal{F}_o$ are likely disjoint except for the proposed extremal solution y^*. It is perhaps a bit unsettling to realize there are infinitely many such subspaces because there are infinitely many choices for the function η or ζ. In any event, each possible variation of y^* must satisfy the condition $\left[\frac{\partial J(\alpha)}{\partial \alpha}\right]_{\alpha=0} = 0$ since we **assume** that y^* is an extreme value of the functional J. It follows that the Equation 9.2 holds for all possible η. Therefore, we ought to expect that any extreme value of the functional $J[y] = \int_{x_1}^{x_2} f(y,y',x)\,dx$ must solve the **Euler Lagrange Equations:**

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] = 0 \quad \text{Euler-Lagrange Equations for} \quad J[y] = \int_{x_1}^{x_2} f(y,y',x)\,dx$$

9.3 variational derivative

The role that η played in the discussion in the preceding section is somewhat similar to the role that the "h" plays in the definition $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$. You might hope we could replace arguments in η with a more direct approach. Physicists have a heuristic way of making such arguments in terms of the variation δ. They would cast the arguments in the last page by just "taking the variation of J". Let me give you their formal argument,

$$\begin{aligned}
\delta J &= \delta\left[\int_{x_1}^{x_2} f(y,y',x)\,dx\right] \\
&= \left[\int_{x_1}^{x_2} \delta f(y,y',x)\,dx\right] \\
&= \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta\left(\tfrac{dy}{dx}\right) + \frac{\partial f}{\partial x}\delta x\right)dx \\
&= \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\frac{d}{dx}(\delta y)\right)dx \\
&= \frac{\partial f}{\partial y'}\delta y\bigg|_{x_1}^{x_2} + \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\right)\delta y\,dx
\end{aligned} \tag{9.3}$$

Therefore, since $\delta y = 0$ at the endpoints of integration, the Euler-Lagrange equations follow from $\delta J = 0$. Now, if you're like me, the argument above is less than satisfying since we never actually defined what it means to "take δ" of something. Also, why could I commute the variational δ and $\frac{d}{dx}$)? That said, the formal method is not without use since it allows the focus to be on the Euler Lagrange equations rather than the technical details of the variation[1].

[1]to be clear, these notes are not a rigorous treatment of variational calculus, the discipline requires more attention to analysis than I offer in this course.

9.4 Euler-Lagrange examples

I present a few standard examples in this section. We make use of the calculation in the last section. We also need the following reformulation of the Euler-Lagrange equation:

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left[f - y'\frac{\partial f}{\partial y'}\right] = 0.$$

This is known as the Beltrami identity, it can be derived by combining the chain-rule $\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}y' + \frac{\partial f}{\partial y'}y''$ and the standard Euler-Lagrange equation $\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]$. If the regular Euler Lagrange equation is troublesome to solve then it may be helpful to solve this alternate version.

9.4.1 shortest distance between two points in plane

If s denotes the arclength in the xy-plane then the pythagorean theorem gives $ds^2 = dx^2 + dy^2$ infinitesimally. Thus, $ds = \sqrt{1 + \frac{dy}{dx}^2}\, dx$ and we may add up all the little distances ds to find the total length between two given points (x_1, y_1) and (x_2, y_2):

$$J[y] = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}\, dx$$

Identify that we have $f(y, y', x) = \sqrt{1 + (y')^2}$. Calculate then,

$$\frac{\partial f}{\partial y} = 0 \qquad \text{and} \qquad \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}.$$

Euler Lagrange equations yield,

$$\frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] = \frac{\partial f}{\partial y} \qquad \Rightarrow \qquad \frac{d}{dx}\left[\frac{y'}{\sqrt{1 + (y')^2}}\right] = 0 \qquad \Rightarrow \qquad \frac{y'}{\sqrt{1 + (y')^2}} = k$$

where $k \in \mathbb{R}$ is constant with respect to x. Moreover, square both sides to reveal

$$\frac{(y')^2}{1 + (y')^2} = k^2 \qquad \Rightarrow \qquad (y')^2 = \frac{k^2}{1 - k^2} \qquad \Rightarrow \qquad \frac{dy}{dx} = \pm\sqrt{\frac{k^2}{1 - k^2}} = m$$

where I have defined m is defined in the obvious way. We find solutions $y = mx + b$. Finally, we can find m, b to fit the given pair of points (x_1, y_1) and (x_2, y_2) as follows:

$$y_1 = mx_1 + b \qquad \text{and} \qquad y_2 = mx_2 + b \qquad \Rightarrow \qquad y = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

provided $x_1 \neq x_2$. If $x_1 \neq x_2$ and $y_1 \neq y_2$ then we could perform the same calculation as above with the roles of x and y interchanged,

$$J[x] = \int_{y_1}^{y_2} \sqrt{1 + (x')^2}\, dy$$

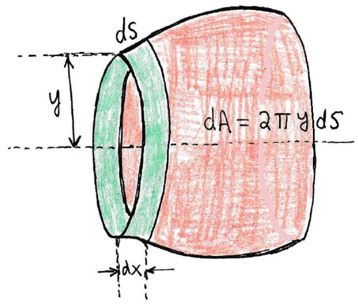
where $x' = dx/dy$ and the Euler Lagrange equations would yield the solution

$$x = x_1 + \frac{x_2 - x_1}{y_2 - y_1}(y - y_1).$$

Finally, if both coordinates are equal then $(x_1, y_1) = (x_2, y_2)$ and the shortest path between these points is the trivial path, the armchair solution. Silly comments aside, we have shown that a straight line provides the curve with the shortest arclength between any two points in the plane.

9.4.2 surface of revolution with minimal area

Suppose we wish to revolve some curve which connects (x_1, y_1) and (x_2, y_2) around the x-axis. A surface constructed in this manner is called a **surface of revolution**. In calculus we learn how to calculate the surface area of such a shape. One can imagine deconstructing the surface into a sequence of ribbons. Each ribbon at position x will have a "radius" of y and a width of dx however, because the shape is tilted the area of the ribbon works out to $dA = 2\pi y ds$ where ds is the arclength. I made a ribbon green in the picture below. You can imagine many ribbons approximating the surface, although, I made no attempt to draw those here:



If we choose x as the parameter this yields $dA = 2\pi y\sqrt{1+(y')^2}\,dx$. To find the surface of minimal surface area we ought to consider the functional:

$$A[y] = \int_{x_1}^{x_2} 2\pi y\sqrt{1+(y')^2}\,dx$$

Identify that $f(y, y', x) = 2\pi y\sqrt{1+(y')^2}$ hence $f_y = 2\pi\sqrt{1+(y')^2}$ and $f_{y'} = 2\pi y y'/\sqrt{1+(y')^2}$. The usual Euler-Lagrange equations are not easy to solve for this problem, instead consider the Beltrami identity version:

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left[f - y'\frac{\partial f}{\partial y'}\right] = 0.$$

Hence,

$$\frac{d}{dx}\left[2\pi y\sqrt{1+(y')^2} - \frac{2\pi y(y')^2}{\sqrt{1+(y')^2}}\right] = 0$$

Dividing by 2π and making a common denominator,

$$\frac{d}{dx}\left[\frac{y}{\sqrt{1+(y')^2}}\right] = 0 \qquad \Rightarrow \qquad \frac{y}{\sqrt{1+(y')^2}} = k$$

where k is a constant with respect to x. Squaring the equation above yields

$$\frac{y^2}{1+(\frac{dy}{dx})^2} = k^2 \qquad \Rightarrow \qquad y^2 - k^2 = k^2(\tfrac{dy}{dx})^2$$

Solve for dx, integrate, assuming the given points are in the first quadrant,

$$x = \int dx = \int \frac{k dy}{\sqrt{y^2-k^2}} = k\cosh^{-1}(\tfrac{y}{k}) + c$$

Hence,

$$\boxed{y = k\cosh\left(\frac{x-c}{k}\right)}$$

generates the surface of revolution of least area between two points. These shapes are called **Catenoids** they can be observed in the formation of soap bubble between rings. There is a vast literature on this subject and there are many cases to consider, I simply exhibit a simple solution. For a given pair of points it is not immediately obvious if there exists a solution to the Euler-Lagrange equations which fits the data. (see page 622 of Arfken).

9.4.3 Braichistochrone

Suppose a particle slides freely along some curve from (x_1, y_1) to $(x_2, y_2) = (0, 0)$ under the influence of gravity where we take y to be the vertical direction. **What is the curve of quickest descent?** Notice that if $x_1 = 0$ then the answer is easy to see, however, if $x_1 \neq 0$ then the question is not trivial. To solve this problem we must first offer a functional which accounts for the time of descent. Note that the speed $v = ds/dt$ so we'd clearly like to minimize $J = \int_{(0,0)}^{(x_1,y_1)} \frac{ds}{v}$. Since the object is assumed to fall freely we may assume that energy is conserved in the motion hence

$$\frac{1}{2}mv^2 = mg(y - y_1) \qquad \Rightarrow \qquad v = \sqrt{2g(y_1 - y)}$$

As we've discussed in previous examples, $ds = \sqrt{1 + (y')^2}dt$ so we find

$$J[y] = \int_0^{x_1} \underbrace{\sqrt{\frac{1 + (y')^2}{2g(y_1 - y)}}}_{f(y,y',x)} \, dx$$

Notice that the modified Euler-Lagrange equations $\frac{\partial f}{\partial x} - \frac{d}{dx}\left[f - y'\frac{\partial f}{\partial y'}\right] = 0$ are convenient since $f_x = 0$. We calculate that

$$\frac{\partial f}{\partial y'} = \frac{1}{2\sqrt{\frac{1+(y')^2}{2g(y_1-y)}}}\frac{2y'}{2g(y_1 - y)} = \frac{y'}{\sqrt{2g(y_1 - y)(1 + (y')^2)}}$$

Hence there should exist some constant $1/(k\sqrt{2g})$ such that

$$\sqrt{\frac{1 + (y')^2}{2g(y_1 - y)}} - \frac{(y')^2}{\sqrt{2g(y_1 - y)(1 + (y')^2)}} = \frac{1}{k\sqrt{2g}}$$

It follows that,

$$\frac{1}{\sqrt{(y_1 - y)(1 + (y')^2)}} = \frac{1}{k} \qquad \Rightarrow \qquad \Big(y_1 - y\Big)\left(1 + \left(\frac{dy}{dx}\right)^2\right) = k^2$$

We need to solve for dy/dx,

$$(y_1 - y)\left(\frac{dy}{dx}\right)^2 = k^2 - y_1 + y \qquad \Rightarrow \qquad \left(\frac{dy}{dx}\right)^2 = \frac{y + k^2 - y_1}{y_1 - y}$$

Or, relabeling constants $a = y_1$ and $b = k^2 - y_1$ and we must solve

$$\frac{dy}{dx} = \pm\sqrt{\frac{b + y}{a - y}} \qquad \Rightarrow \qquad x = \pm\int\sqrt{\frac{a - y}{b + y}}\, dy$$

The integral is not trivial. It turns out that the solution is a cycloid (Arfken p. 624):

$$\boxed{x = \frac{a+b}{2}\Big(\theta + \sin(\theta)\Big) - d \qquad y = \frac{a+b}{2}\Big(1 - \cos(\theta)\Big) - b}$$

This is the curve that is traced out by a point on a wheel as it travels. If you take this solution and calculate $J[y_{cycloid}]$ you can show the time of descent is simply

$$T = \frac{\pi}{2}\sqrt{\frac{y_1}{2g}}$$

if the mass begins to descend from (x_2, y_2). But, this point has no connection with (x_1, y_1) except that they both reside on the same cycloid. It follows that the period of a pendulum that follows a cycloidal path is indpendent of the starting point on the path. This is not true for a circular pendulum in general, we need the small angle approximation to derive simple harmonic motion.
It turns out that it is possible to make a pendulum follow a cycloidal path if you let the string be guided by a frame which is also cycloidal. The neat thing is that even as it loses energy it still follows a cycloidal path and hence has the same period. The "Brachistochrone" problem was posed by Johann Bernoulli in 1696 and it actually predates the variational calculus of Lagrange by some 50 or so years. This problem and ones like it are what eventually prompted Lagrange and Euler to systematically develop the subject. Apparently Galileo also studied this problem however lacked the mathematics to crack it.

9.5 Euler-Lagrange equations for several dependent variables

We still consider problems with just one independent parameter underlying everything. For problems of classical mechanics this is almost always time t. In anticipation of that application we choose to use the usual physics notation in the section. We suppose that our functional depends on functions $y_1, y_2, \dots, y_n$ of time t along with their time derivatives $\dot{y}_1, \dot{y}_2, \dots, \dot{y}_n$. We again suppose the functional of interest is an integral of a **Lagrangian** function f from time t_1 to time t_2,

$$J[(y_i)] = \int_{t_1}^{t_2} f(y_i, \dot{y}_i, t)\, dt$$

here we use (y_i) as shorthand for $(y_1, y_2, \dots, y_n)$ and $(\dot{y}_i)$ as shorthand for $(\dot{y}_1, \dot{y}_2, \dots, \dot{y}_n)$. We suppose that n-conditions are given for each of the endpoints in this problem; $y_i(t_1) = y_{i1}$ and $y_i(t_2) = y_{i2}$. Moreover, we define $\mathcal{F}_o$ to be the set of paths from $\mathbb{R}$ to $\mathbb{R}^n$ subject to the conditions just stated. We now set out to find necessary conditions on a proposed solution to the extreme value problem for the functional J above. As before let's assume that an extremal solution $y* \in \mathcal{F}_o$ exists. Moreover, imagine varying the solution by some variational function $\eta = (\eta_i)$ which has $\eta(t_1) = (0, 0, \dots, 0)$ and $\eta(t_2) = (0, 0, \dots, 0)$. Consequently the family of paths defined below are all in $\mathcal{F}_o$,

$$y(t, \alpha) = y^*(t) + \alpha\eta(t)$$

Thus $y(t, 0) = y^*$. In terms of component functions we have that

$$y_i(t, \alpha) = y_i^*(t) + \alpha\eta_i(t).$$

You can identify that $\delta y_i = y_i(t,\alpha) - y_i^*(t) = \alpha\eta_i(t)$. Since y^* is an extreme solution we should expect that $\left(\frac{\partial J}{\partial \alpha}\right)_{\alpha=0} = 0$. Differentiate the functional with respect to α and make use of the chain rule for f which is a function of some $2n+1$ variables,

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[\int_{t_1}^{t_2} f(y_i(t,\alpha), \dot{y}_i(t,\alpha), t)\, dt\right] \\
&= \int_{t_1}^{t_2} \sum_{j=1}^{n}\left(\frac{\partial f}{\partial y_j}\frac{\partial y_j}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}_j}\frac{\partial \dot{y}_j}{\partial \alpha}\right) dt \\
&= \int_{t_1}^{t_2} \sum_{j=1}^{n}\left(\frac{\partial f}{\partial y_j}\eta_j + \frac{\partial f}{\partial \dot{y}_j}\frac{d\eta_j}{dt}\right) dt \\
&= \sum_{j=1}^{n} \frac{\partial f}{\partial \dot{y}_j}\eta\Big|_{t_1}^{t_2} + \int_{t_1}^{t_2}\sum_{j=1}^{n}\left(\frac{\partial f}{\partial y_j} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}_j}\right)\eta_j\, dt
\end{aligned} \tag{9.4}$$

Since $\eta(t_1) = \eta(t_2) = 0$ the first term vanishes. Moreover, since we may repeat this calculation for all possible variations about the optimal solution y^* it follows that we obtain a set of Euler-Lagrange equations for each component function of the solution:

$$\boxed{\frac{\partial f}{\partial y_j} - \frac{d}{dt}\left[\frac{\partial f}{\partial \dot{y}_j}\right] = 0 \quad j = 1, 2, \dots n \quad \text{Euler-Lagrange Eqns. for} \quad J[(y_i)] = \int_{t_1}^{t_2} f(y_i, \dot{y}_i, t)\, dt}$$

Often we simply use $y_1 = x$, $y_2 = y$ and $y_3 = z$ which denote the position of particle or perhaps just the component functions of a path which gives the geodesic on some surface. In either case we should have 3 sets of Euler-Lagrange equations, one for each coordinate. We will also use non-Cartesian coordinates to describe certain Lagrangians. We develop many useful results for set-up of Lagrangians in non-Cartesian coordinates in the next section.

9.5.1 free particle Lagrangian

For a particle of mass m the kinetic energy K is given in terms of the time derivatives of the coordinate functions x, y, z as follows:

$$K = \tfrac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)$$

Construct a functional by integrating the kinetic energy over time t,

$$S = \int_{t_1}^{t_2} \tfrac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)\, dt$$

The Euler-Lagrange equations for this functional are

$$\frac{\partial K}{\partial x} = \frac{d}{dt}\left[\frac{\partial K}{\partial \dot{x}}\right] \qquad \frac{\partial K}{\partial y} = \frac{d}{dt}\left[\frac{\partial K}{\partial \dot{y}}\right] \qquad \frac{\partial K}{\partial z} = \frac{d}{dt}\left[\frac{\partial K}{\partial \dot{z}}\right]$$

Since $\frac{\partial K}{\partial \dot{x}} = m\dot{x}$, $\frac{\partial K}{\partial \dot{y}} = m\dot{y}$ and $\frac{\partial K}{\partial \dot{z}} = m\dot{z}$ it follows that

$$\boxed{0 = m\ddot{x} \qquad 0 = m\ddot{y} \qquad 0 = m\ddot{z}.}$$

You should recognize these as Newton's equation for a particle with no force applied. The solution is $(x(t), y(t), z(t)) = (x_o + tv_x, y_o + tv_y, z_o + tv_z)$ which is uniform rectilinear motion at constant velocity (v_x, v_y, v_z). The solution to Newton's equation minimizes the integral of the Kinetic energy. Generally the quantity S is called the **action** and Hamilton's Principle states that the laws of physics all arise from minimizing the action of the physical phenomena. We'll return to this discussion in a later section.

9.5.2 path of least distance between points in $\mathbb{R}^3$

The arclength integral from $p = 0$ to $q = (q_x, q_y, q_z)$ in $\mathbb{R}^3$ is most naturally given from the parametric viewpoint:

$$S = \int_0^1 \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\ dt$$

We assume $(x(0), y(0), z(0)) = (0, 0, 0)$ and $(x(1), y(1), z(1)) = q$ and it should be clear that the integral above calculates the arclength. The Euler-Lagrange equations for x, y, z are

$$\frac{d}{dt}\left[\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right] = 0, \qquad \frac{d}{dt}\left[\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right] = 0, \qquad \frac{d}{dt}\left[\frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right] = 0.$$

It follows that there exist constants, say a, b and c, such that

$$a = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}, \qquad b = \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}, \qquad c = \frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}.$$

These equations are said to be **coupled** since each involves derivatives of the others. We usually need a way to uncouple the equations if we are to be successful in solving the system. We can calculate, and equate each with the constant 1:

$$1 = \frac{\dot{x}}{a\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \frac{\dot{y}}{b\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \frac{\dot{z}}{c\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}.$$

But, multiplying by the denominator reveals an interesting identity

$$\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} = \frac{\dot{x}}{a} = \frac{\dot{y}}{b} = \frac{\dot{z}}{c}$$

The solution has the form, $x(t) = tq_x$, $y(t) = tq_y$ and $z(t) = tq_z$. Therefore,

$$(x(t), y(t), z(t)) = t(q_x, q_y, q_z) = tq.$$

for $0 \leq t \leq 1$. These are the parametric equations for the line segment from the origin to q.

9.6 geodesics with respect to the Euclidean metric

A **geodesic** is a path of least length on a space. It turns out there are generally many different ways to formulate the length of a path on a given space, so, to be clear, we study the standard Euclidean path length which is governed by the Euclidean metric $ds^2 = dx_1^2 + dx_2^2 + \cdots + dx_n^2$. The square root in the functional of the last subsection certainly complicated the calculation. It is intuitively clear that if we add up squared line elements ds^2 to give a minimum then that ought to correspond to the minimum for the sum of the positive square roots ds of those elements. Let's check if my conjecture works for $\mathbb{R}^3$:

$$S = \int_0^1 (\underbrace{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}_{f(x,y,z,\dot{x},\dot{y},\dot{z})})\ dt$$

This gives us the Euler Lagrange equations below:

$$\ddot{x} = 0, \qquad \ddot{y} = 0, \qquad \ddot{z} = 0$$

The solution of these equations is clearly a line. In this formalism the equations were uncoupled from the outset.

Definition 9.6.1.

The Euclidean metric of $\mathbb{R}^3$ is $ds^2 = dx^2 + dy^2 + dz^2$. Generally, for orthogonal curvelinear coordinates u, v, w we calculate $ds^2 = \frac{1}{||\nabla u||^2} du^2 + \frac{1}{||\nabla v||^2} dv^2 + \frac{1}{||\nabla w||^2} dw^2$. We use this as a guide for constructing functionals which calculate arclength or speed. For the plane, $ds^2 = dx^2 + dy^2$.

The beauty of the metric is that it allows us to calculate in other coordinates, consider

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

For which we have implicit inverse coordinate transformations $r^2 = x^2 + y^2$ and $\theta = \tan^{-1}(y/x)$. From these inverse formulas we calculate:

$$\nabla r \; = < x/r, y/r > \qquad \nabla\theta \; = < -y/r^2, x/r^2 >$$

Thus, $||\nabla r|| = 1$ whereas $||\nabla\theta|| = 1/r$. We find that the metric in polar coordinates takes the form:

$$\boxed{ds^2 = dr^2 + r^2 d\theta^2}$$

Physicists and engineers tend to like to think of these as arising from calculating the length of infinitesimal displacements in the r or θ directions. Generically, for u, v, w coordinates

$$dl_u = \frac{1}{||\nabla u||} du \qquad dl_v = \frac{1}{||\nabla v||} dv \qquad dl_w = \frac{1}{||\nabla w||} dw$$

and $ds^2 = dl_u^2 + dl_v^2 + dl_w^2$. So in that notation we just found $dl_r = dr$ and $dl_\theta = r d\theta$. Notice then that cylindircal coordinates have the metric,

$$\boxed{ds^2 = dr^2 + r^2 d\theta^2 + dz^2.}$$

For spherical coordinates $x = r\cos(\phi)\sin(\theta)$, $y = r\sin(\phi)\sin(\theta)$ and $z = r\cos(\theta)$ (here $0 \leq \phi \leq 2\pi$ and $0 \leq \theta \leq \pi$, physics notation). Calculation of the metric follows from the line elements,

$$dl_r = dr \qquad dl_\phi = r\sin(\theta) d\phi \qquad dl_\theta = r d\theta$$

Thus,

$$\boxed{ds^2 = dr^2 + r^2\sin^2(\theta) d\phi^2 + r^2 d\theta^2.}$$

We now have all the tools we need for examples in spherical or cylindrical coordinates. What about other cases? In general, given some p-manifold in $\mathbb{R}^n$ how does one find the metric on that manifold? If we are to follow the approach of this section we'll need to find coordinates on $\mathbb{R}^n$ such that the manifold S is described by setting all but p of the coordinates to a constant. For example, in $\mathbb{R}^4$ we have generalized cylindircal coordinates (r, ϕ, z, t) defined implicitly by the equations below

$$x = r\cos(\phi), \qquad y = r\sin(\phi), \qquad z = z, \qquad t = t$$

On the hyper-cylinder $r = R$ we have the metric $ds^2 = R^2 d\theta^2 + dz^2 + dw^2$. There are mathematicians/physicists whose careers are founded upon the discovery of a metric for some manifold. This is generally a difficult task.

9.6.1 geodesic on cylinder

The equation of a cylinder of radius R is most easily framed in cylindrical coordinates (r, θ, z); the equation is merely $r = R$ hence the metric reads

$$ds^2 = R^2 d\theta^2 + dz^2$$

Therefore, we ought to minimize the following functional in order to locate the parametric equations of a geodesic on the cylinder: note $ds^2 = \left(R^2 \frac{d\theta^2}{dt^2} + \frac{dz^2}{dt^2}\right)dt^2$ thus:

$$S = \int (R^2\dot{\theta}^2 + \dot{z}^2)\, dt$$

Euler-Lagrange equations for the dependent variables θ and z are simply:

$$\ddot{\theta} = 0 \qquad \ddot{z} = 0.$$

We can integrate twice to find solutions

$$\boxed{\theta(t) = \theta_o + At \qquad z(t) = z_o + Bt}$$

Therefore, the geodesic on a cylinder is simply the line connecting two points in the plane which is curved to assemble the cylinder. Simple cases that are easy to understand:

1. Geodesic from $(R\cos(\theta_o), R\sin(\theta_o), z_1)$ to $(R\cos(\theta_o), R\sin(\theta_o), z_2)$ is parametrized by $\theta(t) = \theta_o$ and $z(t) = z_1 + t(z_2 - z_1)$ for $0 \leq t \leq 1$. Technically, there is some ambiguity here since I never declared over what range the t is to range. Could pick other intervals, we could use z at the parameter is we wished then $\theta(z) = \theta_o$ and $z = z$ for $z_1 \leq z \leq z_2$

2. Geodesic from $(R\cos(\theta_1), R\sin(\theta_1), z_o)$ to $(R\cos(\theta_2), R\sin(\theta_2), z_o)$ is parametrized by $\theta(t) = \theta_1 + t(\theta_2 - \theta_1)$ and $z(t) = z_o$ for $0 \leq t \leq 1$.

3. Geodesic from $(R\cos(\theta_1), R\sin(\theta_1), z_1)$ to $(R\cos(\theta_2), R\sin(\theta_2), z_2)$ is parametrized by

$$\theta(t) = \theta_1 + t(\theta_2 - \theta_1) \qquad z(t) = z_1 + t(z_2 - z_1)$$

   You can eliminate t and find the equation $z = \frac{z_2 - z_1}{\theta_2 - \theta_1}(\theta - \theta_1)$ which again just goes to show you this is a line in the curved coordinates.

9.6.2 geodesic on sphere

The equation of a sphere of radius R is most easily framed in spherical coordinates (r, ϕ, θ) where $x = r\cos\phi\sin\theta, y = r\sin\phi\sin\theta, z = r\cos\theta$; the equation is merely $r = R$ hence the metric reads

$$ds^2 = R^2\sin^2(\theta)d\phi^2 + R^2 d\theta^2.$$

Therefore, we ought to minimize the following functional in order to locate the parametric equations of a geodesic on the sphere: note $ds^2 = \left(R^2\sin^2(\theta)\frac{d\phi^2}{dt^2} + R^2\frac{d\theta^2}{dt^2}\right)dt^2$ thus, dividing by R^2 to reduce clutter, we find:

$$S = \int (\underbrace{\sin^2(\theta)\dot{\phi}^2 + \dot{\theta}^2}_{f(\theta,\phi,\dot{\theta},\dot{\phi})})\, dt$$

Euler-Lagrange equations for the dependent variables ϕ and θ are simply: $f_\theta = \frac{d}{dt}(f_{\dot\theta})$ and $f_\phi = \frac{d}{dt}(f_{\dot\phi})$ which yield:

$$2\sin(\theta)\cos(\theta)\dot\phi^2 = \tfrac{d}{dt}(2\dot\theta) \qquad 0 = \frac{d}{dt}\Big(2\sin^2(\theta)\dot\phi\Big).$$

The intersection of the sphere $r = R$ and the $z = 0$ plane has $\theta = \pi/2$. Notice $\sin(\theta) = 0$ in this case hence the Euler-Lagrange equations reduce to $0 = \ddot\theta$ which has solution $\theta = c_1 t + c_2$. In other words, the equatorial circle forms a geodesic of the sphere. Hence, by symmetry of the xy-plane with every other plane through the origin we find all great circles of the sphere $r = R$ are geodesics[2].

9.6.3 geodesic on graph

Let's consider the graph $z = f(x, y)$ then $\dot z = f_x\dot x + f_y\dot y$ thus the metric on the graph is given in terms of x and y coordinates and we form the action by:

$$S = \int \Big(\dot x^2 + \dot y^2 + (f_x\dot x + f_y\dot y)^2\Big)\, dt$$

Notice the Lagrangian $L = \dot x^2 + \dot y^2 + (f_x\dot x + f_y\dot y)^2$ hence the Euler-Lagrange equations

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\left[\frac{\partial L}{\partial \dot x}\right] \qquad \& \qquad \frac{\partial L}{\partial y} = \frac{d}{dt}\left[\frac{\partial L}{\partial \dot y}\right].$$

Consider $f(x, y) = x^2 + y^2$, that is $z = x^2 + y^2$ which is a paraboloid. Then $f_x = 2x$ and $f_y = 2y$ and

$$L = \dot x^2 + \dot y^2 + (2x\dot x + 2y\dot y)^2.$$

Thus the Euler-Lagrange equations are:

$$4\dot x\,(2x\dot x + 2y\dot y) = \frac{d}{dt}\Big[2\dot x + 4x\,(2x\dot x + 2y\dot y)\Big] \;\;\&\;\; 4\dot y\,(2x\dot x + 2y\dot y) = \frac{d}{dt}\Big[2\dot y + 4y\,(2x\dot x + 2y\dot y)\Big].$$

I bet this is easier in cylindrical coordinates, $z = r^2$ hence $dz = 2rdr$ and we find

$$ds^2 = dr^2 + r^2d\theta^2 + dz^2 = (1 + 4r^2)dr^2 + r^2d\theta^2$$

Then $L = (1 + 4r^2)\dot r^2 + r^2\dot\theta^2$ and the Euler-Lagrange equations for r and θ are as follows:

$$\frac{\partial L}{\partial r} = \frac{d}{dt}\left[\frac{\partial L}{\partial \dot r}\right] \qquad \& \qquad \frac{\partial L}{\partial \theta} = \frac{d}{dt}\left[\frac{\partial L}{\partial \dot\theta}\right].$$

Thus,

$$8r\dot r^2 + 2r\dot\theta^2 = \frac{d}{dt}\Big[2\dot r(1 + 4r^2)\Big] \qquad \& \qquad 0 = \frac{d}{dt}\Big[2r^2\dot\theta\Big]$$

You might think a horizontal circle on the paraboloid is a geodesic. That is $r = R$ with $\dot r = 0$. You can check that such a circle does not solve the Euler-Lagrange equations above and hence the intersection of a paraboloid with a horizontal plane does not form a geodesic.

[2] I have not shown these are the only solutions of the Euler-Lagrange equations, I leave that to the curious reader.

9.7 Lagrangian mechanics

Lagrangian mechanics allows us to formulate problems of classical mechanics in terms of **generalized coordinates**. To formulate the Lagrangian for a problem you need to identify a suitable set of coordinates then derive the potential and kinetic energy functions to which they correspond. There is no particular need to use Cartesian coordinates. Other systems like cylindrical or spherical or far less familar problem-specific systems are often used. For example, the double-pendulum is naturally solved using the angle off the vertical for both pivot points. I will probably set that up in class, it is not currently in these notes. Also, there is a simple procedure to implement constraints in terms of a Lagrange multiplier term for the Lagrangian. That formalism naturally allows us to derive forces necessary to impose certain geometric constraints. Once more, I have not typed up that material, but time permitting I will demonstrate an example.

9.7.1 basic equations of classical mechanics summarized

Classical mechanics is the study of massive particles at relatively low velocities. Let me refresh your memory about the basics equations of Newtonian mechanics. Our goal in this section will be to rephrase Newtonian mechanics in the variational langauge and then to solve problems with the Euler-Lagrange equations. Newton's equations tell us how a particle of mass m evolves through time according to the net-force impressed on m. In particular,

$$m\frac{d^2\vec{r}}{dt^2} = \vec{F}$$

If m is not constant then you may recall that it is better to use momentum $\vec{P} = m\vec{v} = m\frac{d\vec{r}}{dt}$ to set-up Newton's 2nd Law:

$$\frac{d\vec{P}}{dt} = \vec{F}$$

In terms of components we have a system of differential equations with indpendent variable time t. If we use position as the dependent variable then Newton's 2nd Law gives three second order ODEs,

$$m\ddot{x} = F_x \qquad m\ddot{y} = F_y \qquad m\ddot{z} = F_z$$

where $\vec{r} = (x, y, z)$ and the dots denote time-derivatives. Moreover, $\vec{F} =< F_x, F_y, F_z >$ is the sum of the forces that act on m. In contrast, if you work with momentum then you would want to solve six first order ODEs,

$$\dot{P}_x = F_x \qquad \dot{P}_y = F_y \qquad \dot{P}_z = F_z$$

and $P_x = m\dot{x}$, $P_y = m\dot{y}$ and $P_z = m\dot{z}$. These equations are easiest to solve when the force is not a function of velocity or time. In particular, if the force $\vec{F}$ is conservative then there exists a potential energy function $U : \mathbb{R}^3 \to \mathbb{R}$ such that $\vec{F} = -\nabla U$. We can prove that in the case the force is conservative the total energy is conserved.

9.7.2 kinetic and potential energy, formulating the Lagrangian

Recall the kinetic energy is $T = \frac{1}{2}m||\vec{v}||^2$, in Cartesian coordinates this gives us the formula:

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

If $\vec{F}$ is a conservative force then it is independent of path so we may construct the potential energy function as follows:

$$U(\vec{r}) = -\int_{\mathcal{O}}^{\vec{r}} \vec{F} \cdot d\vec{r}$$

Here $\mathcal{O}$ is the origin for the potential and we can prove that the potential energy constructed in this manner has $\vec{F} = -\nabla U$. We can prove that the total (mechanical) energy $E = T + U$ for a conservative system is a constant; $dE/dt = 0$. Hopefully these comments are at least vaguely familiar from some physics course in your distant memory. If not relax, calculationally this chapter is self-contained, read onward.

We already calculated that if we use T as the Lagrangian then the Euler-Lagrange equations produce Newton's equations in the case that the force is zero (see 9.5.1). Suppose that we define the Lagrangian to be $L = T - U$ for a system governed by a conservative force with potential energy function U. We seek to prove the Euler-Lagrange equations are precisely Newton's equations for this conservative system[3] Generically we have a Lagrangian of the form

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - U(x, y, z).$$

We wish to find extrema for the functional $S = \int L(t)\, dt$. This yields three sets of Euler-Lagrange equations, one for each dependent variable x, y or z

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{x}}\right] = \frac{\partial L}{\partial x} \qquad \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{y}}\right] = \frac{\partial L}{\partial y} \qquad \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{z}}\right] = \frac{\partial L}{\partial z}.$$

Note that $\frac{\partial L}{\partial \dot{x}} = m\dot{x}$, $\frac{\partial L}{\partial \dot{y}} = m\dot{y}$ and $\frac{\partial L}{\partial \dot{z}} = m\dot{z}$. Also note that $\frac{\partial L}{\partial x} = -\frac{\partial U}{\partial x} = F_x$, $\frac{\partial L}{\partial y} = -\frac{\partial U}{\partial y} = F_y$ and $\frac{\partial L}{\partial z} = -\frac{\partial U}{\partial z} = F_z$. It follows that

$$\boxed{m\ddot{x} = F_x \qquad m\ddot{y} = F_y \qquad m\ddot{z} = F_z.}$$

Of course this is precisely $m\vec{a} = \vec{F}$ for a net-force $\vec{F} = < F_x, F_y, F_z >$. We have shown that **Hamilton's principle** reproduces Newton's Second Law for conservative forces. Let me take a moment to state it.

Definition 9.7.1. Hamilton's Principle:

If a physical system has generalized coordinates q_j with velocities $\dot{q}_j$ and Lagrangian $L = T - U$ then the solutions of physics will minimize the action S defined below:

$$\boxed{S = \int_{t_1}^{t_2} L(q_j, \dot{q}_j, t)\, dt}$$

Mathematically, this means the variation $\delta S = 0$ for physical trajectories.

This is a necessary condition for solutions of the equations of physics. Sufficient conditions are known, you can look in any good variational calculus text. You'll find analogues to the second derivative test for variational differentiation. As far as I can tell physicists don't care about this logical gap, probably because the solutions to the Euler-Lagrange equations are the ones for which they are looking.

[3] don't mistake this example as an admission that Lagrangian mechanics is limited to conservative systems. Quite the contrary, Lagrangian mechanics is actually more general than the orginal framework of Newton!

9.7.3 easy physics examples

Now, you might just see this whole exercise as some needless multiplication of notation and formalism. After all, I just told you we just get Newton's equations back from the Euler-Lagrange equations. To my taste the impressive thing about Lagrangian mechanics is that you get to start the problem with energy. Moreover, the Lagrangian formalism handles non-Cartesian coordinates with ease. If you search your memory from classical mechanics you'll notice that you either do constant acceleration, circular motion or motion along a line. What if you had a particle constrained to move in some frictionless ellipsoidal bowl. Or what if you had a pendulum hanging off another pendulum? How would you even write Newtons' equations for such systems? In contrast, the problem is at least easy to set-up in the Lagrangian approach. Of course, solutions may be less easy to obtain.

Example 9.7.2. Projectile motion: *take z as the vertical direction and suppose a bullet is fired with initial velocity $v_o =< v_{ox}, v_{oy}, v_{oz} >$. The potential energy due to gravity is simply $U = mgz$ and kinetic energy is given by $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$. Thus,*

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz$$

Euler-Lagrange equations are simply:

$$\frac{d}{dt}\Big[m\dot{x}\Big] = 0 \qquad \frac{d}{dt}\Big[m\dot{y}\Big] = 0 \qquad \frac{d}{dt}\Big[m\dot{z}\Big] = \frac{\partial}{\partial z}(-mgz) = -mg.$$

Integrating twice and applying initial conditions gives us the (possibly familiar) equations

$$x(t) = x_o + v_{ox}t, \qquad y(t) = y_o + v_{oy}t, \qquad z(t) = z_o + v_{oz}t - \tfrac{1}{2}gt^2.$$

Example 9.7.3. Simple Pendulum: *let θ denote angle measured off the vertical for a simple pendulum of mass m and length l. Trigonmetry tells us that*

$$x = l\sin(\theta) \qquad y = l\cos(\theta) \qquad \Rightarrow \qquad \dot{x} = l\cos(\theta)\dot{\theta} \qquad y = -l\sin(\theta)\dot{\theta}$$

Thus $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}ml^2\dot{\theta}^2$. Also, the potential energy due to gravity is $U = -mgl\cos(\theta)$ which gives us

$$L = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos(\theta)$$

Then, the Euler-Lagrange equation in θ is simply:

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{\theta}}\right] = \frac{\partial L}{\partial \theta} \qquad \Rightarrow \qquad \frac{d}{dt}(ml^2\dot{\theta}) = -mgl\sin(\theta) \qquad \Rightarrow \qquad \ddot{\theta} + \frac{g}{l}\sin(\theta) = 0.$$

In the small angle approximation, $\sin(\theta) = \theta$ then we have the solution $\theta(t) = \theta_o\cos(\omega t + \phi_o)$ for angular frequency $\omega = \sqrt{g/l}$

Remark 9.7.4.

I intend to show www.supermath.info/centralforcesolution.pdf in lecture where I show how to solve the central force problem. That discussion brings us to an explicit solution of Kepler's Laws which derives the explicit elliptical form of the orbits.