DIFFERENTIAL EQUATIONS
RECITATION NOTES WITHOUT SOLUTIONS

James Holland

math.rutgers.edu/~jch258

jch258@math.rutgers.edu

Start Date: 2018-09-03
End Date: 2018-12-23

Updated: 2020-05-10

Contents

Chapter I. Background

Section 1. *Terminology*

It's very important to know some of the common words used in talking about mathematical problems, and equations in particular. Many students think of an equation merely as something to be solved, or the start of a pre-determined procedure. More than a crutch to get through the class, this conception fundamentally cripples their ability to learn more advanced topics.

To start at the very basics, an equation is a statement. Like any statement, it can be either true or false. For example,

1. the equation $y = 5$ is true when y is 5, but is false when y is 4;
2. the equation $2y + 1 = 1$ is only true when $y = 0$;
3. the equation $2y + y = 3y$ is true for all y;
4. the equation $(y - 1)(y + 2) = 0$ is true for $y = 1$ and for $y = -2$;
5. the equation $y + 1 = y$ is never true.

To say that something *solves* an equation means that it makes the equation true. Similarly, something *satisfies* some condition or statement if that statement is true of it. Something *witnesses* that a statement is true, if that statement is true of it. These are all different words for the same thing. A *solution* is just something that solves the equation. So, returning to (4) above, we might say that $y = 1$ is *a* solution to the equation, or that $y = -2$ *witnesses* that $(y - 1)(y + 2) = 0$. We might say that (5) above has *no* solution.

If we can describe all the possible solutions, this description is the *general solution*. Otherwise, we might only know a couple solutions, known as *particular solutions*. For example, $7x^3 + 2x^2 - 100x = 0$ has a particular solution $x = 0$, and the general solution is hard to find. $(x - 1)(x + 1) = 0$ has a general solution of $x = \pm 1$, where $x = 1$ and $x = -1$ are both particular solutions.

So solving an equation is just finding out when that equation is true. This is done through logic and reasoning. Many students, however, are not very good at this, and fail to understand that mathematics is about giving an argument, just as intelligible and legible as any other. This means complete sentences, and giving reasons. For example, consider the following line of flawed reasoning that attempts to show why $2y + y = 3y$ is always true:

1. $2y + y = 3y$.
2. Factor out a y: $(2 + 1)y = 3y$.
3. Divide by y: $2 + 1 = 3$.
4. Simplify: $3 = 3$ ✓

Such reasoning is mind-bogglingly wrong. Consider an analogous line of reasoning in plain English: I'm trying to show that all horses are animals.

1. All horses are animals.
2. For any horse, it is an animal.
3. Since all horses are animals, an animal is an animal, which we know is true.

Clearly this doesn't show that all horses are animals. The issue is that it assumes at the beginning what it wants to show, and then concludes what we already knew. A proper line of reasoning would be the following

1. $3 = 3$ is clear.
2. $2 + 1 = 3$ is also clear.

3. Thus $3y$ is the same as $(2+1)y$, since 3 and $2+1$ are the same thing.
4. By distributing, this means $3y$ and $2y+y$ are the same.
5. Thus $2y+y=3y$.

Section 2. *Calculus Review*

There are two fundamental rules for calculating derivatives: the chain rule , and the product rule . As a kind of inverse operation, these correspond to two rules of integration: u-substitution , and integration by parts respectively.

§ 2 A. Differentiation

In most circumstances, differentiation relies on two rules and a few memorized derivatives.

2 A • 1. Result (the chain rule)

Let f and g be two differentiable functions, each with one input. Therefore the composition $h(x) := f(g(x))$ has a derivative h' given by
$$h'(x) = f'(g(x)) \cdot g'(x).$$

There is a useful mnemonic to remember this result:
$$\frac{\mathrm{d}}{\mathrm{d}x} f(g) = \frac{\mathrm{d}f}{\mathrm{d}g} \frac{\mathrm{d}g}{\mathrm{d}x},$$
where the $\mathrm{d}g$s "cancel out". This mnemonic isn't super helpful in practice, however, as it doesn't tell you where these functions are evaluated. More properly, it should be written as in the chain rule.

Along with the chain rule is the product rule.

2 A • 2. Result (the product rule)

Let f and g be two differentiable functions, each with one input. Therefore the product $h(x) := f(x) \cdot g(x)$ has a derivative h' given by
$$h'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

Both of these results should be very familiar to you, and they allow you to calculate the derivatives of complicated expressions just through syntactical manipulation. Often this happens without any actual thought required beyond remembering a few derivatives. In particular, the following will be important.

2 A • 3. Result

You should know
1. $\frac{\mathrm{d}}{\mathrm{d}x} e^x = e^x$, i.e. e^x isn't changed by differentiation;
2. $\frac{\mathrm{d}}{\mathrm{d}x} x^n = nx^{n-1}$ for any fixed number n like 3, π, $\sqrt{1/2}$, etc.;
3. $\frac{\mathrm{d}}{\mathrm{d}x} \ln|x| = 1/x$;
4. $\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x$; and
5. $\frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x$.

§ 2 B. Integration

Anyone who has taken calculus knows that Integration is much harder than differentiation. In practice, integration relies on two major rules, which are just the reverse of the differentiation rules: u-substitution , and integration by parts. The easier of the two methods is substitution.

2 B • 1. Result (u-substitution)

Let f and u be smooth functions, each with one input. Therefore
$$\int_a^b f(u(x)) \cdot u'(x)\, dx = \int_{u(a)}^{u(b)} f(u)\, du.$$

This again helps show where exactly these things are evaluated, so that if we're instead taking the indefinite integral, it's not hard to see that $\int f(u(x)) \cdot u'(x)\, dx$ is just $F(u(x))$ where $F = \int f\, dx$.

The more difficult to master method of integration is *integration by parts*, a kind of inverse of the product rule. This is difficult partly because the form is less intuitive, but also because it requires more choice and calculation. Often one will need to do several attempts just to get the right choice of functions that allows the integral to be done easily. Occasionally, some further algebra is needed to get the value.

2 B • 2. Result (integration by parts)

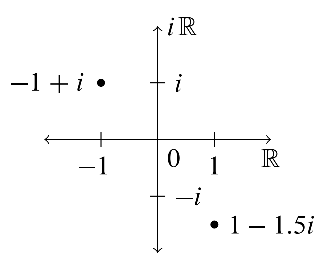
Let f and g be smooth functions, each with one input. Therefore
$$\int_a^b f'(x) \cdot g(x)\, dx = f(x) \cdot g(x)\Big]_a^b - \int_a^b f(x) \cdot g'(x)\, dx.$$

As we can see above, there's no change from where we're evaluating the functions, so there is then no harm in stating the indefinite case as just
$$\int f'g\, dx = fg - \int fg'\, dx.$$
The difficulty then comes in choosing which function one wants to be f' as opposed to g. This method requires that intuition, but also the ability to calculate the derivatives of these functions as well as their integrals, and then integrating their product.

Problems like these are the bread and butter of integration exercises. But often, integrals can be calculated in a slightly more direct way using complex numbers. Not all the time, of course, but in certain cases as we will see.

Section 3. *Complex Numbers Review*

If the real numbers can be viewed as a line, the complex numbers can be viewed as a plane:



This means any complex number z is written as $z = a + bi$, where a is referred to as the *real part* of z, and b is the *imaginary part* of z. In notation,
$$z = \text{Re}(z) + i \cdot \text{Im}(z).$$

And as with any other coordinate plane, we can represent any two points by cartesian coordinates, or by polar coordinates. For example, $-1 + i$ is written in a kind of cartesian form: the sum of its real (-1) and imaginary (1) parts: $p = (-1) + 1 \cdot i$. Alternatively, we can view any point p in $\mathbb{C}$ by its angle with the (positive) real axis, and distance from 0.

Complex numbers can be brought about and defined in many ways: as a vector space, as an algebraic closure, as a quotient, and so on. The value i can be viewed as being anywhere from merely a formal symbol with no inherent connection to the real world to being a fundamental aspect in the algebra of even the real numbers. For our purposes, we have the following defintions.

3 • 1. Definition

Extend the real numbers $\mathbb{R}$ with a value i that satisfies $i^2 = -1$.
An *imaginary number* is a quantity ai for a real number a.
A *complex number* is a quantity $a + bi$ for real numbers a and b. The set of complex numbers is $\mathbb{C}$.
For $z = a + bi$ a complex number,

- the *real part* of z is $\operatorname{Re}(z) = a$.
- The *imaginary part* of z is $\operatorname{Im}(z) = b$.
- The *absolute value* or *magnitude* of z is $|z| = \sqrt{a^2 + b^2}$, a real number.
- The *conjugate* of z is $\overline{z} = a - bi$.

The *(multiplicative) inverse* of a complex number z is a complex number s where $s \cdot z = 1$. If such an s exists, it's called $1/z$.

Note that since $s \cdot 0 = 0 \neq 1$, this means 0 has no multiplicative inverse—i.e. we can't divide by 0. Writing $z = a + bi$ yields that

$$\frac{1}{z} = \frac{1}{a+bi} \cdot \frac{a-bi}{a-bi} = \frac{a-bi}{a^2 - b^2 i^2} = \frac{a-bi}{a^2+b^2} = \frac{\overline{z}}{|z|^2}.$$

Instead of the cartesian form $a + bi$, we can also represent complex numbers in polar form: $re^{i\theta}$. Proving this requires working with whatever definition of e one prefers. So when working with sines and cosines, Euler's[i] formula is a must-have.

3 • 2. Result (Euler's formula)

For any θ, $e^{i\theta} = \cos\theta + i\sin\theta$.

The usual exponentiation laws still hold: $e^z e^s = e^{z+s}$. And so when n is an integer, $(e^z)^n = e^{z \cdot n}$.

But note that logarithms now require a choice, e.g. $e^{\pi i} = -i = e^{3\pi i}$ so we must choose whether $\ln i$ is πi or $3\pi i$ (or any of the other odd multiples of πi). Similarly, taking roots requires more choice than before, e.g. $1^4 = (-1)^4 = i^4 = (-i)^4 = 1$, so we must be careful in taking the fourth root of 1, for example. As it turns out, such a choice is rarely important for our purposes, just as it is in the real case or in functions like $\tan^{-1}$.

Returning to Euler's formula, this implies $|e^{i}| = 1$ for all θ. Furthermore, this gives a quick way to verify some trigonometric equalities, e.g.

$$\begin{aligned}(e^{i\theta})^2 = (\cos\theta + i\sin\theta)^2 &= \cos^2\theta - \sin^2\theta + i2\sin\theta\cos\theta \\ &= e^{i2\theta} = \cos(2\theta) + i\sin(2\theta).\end{aligned}$$

Equating the real and imaginary parts yields that $\cos(2\theta) = \cos^2\theta - \sin^2\theta$, and $\sin(2\theta) = 2\sin\theta\cos\theta$.

The central question when a new concept is brought up like this is "what good is it?". The answer for our purposes is mostly the ease of calculating integrals. This is accomplished by reducing the many hard-to-remember trigonometric identities to easy-to-confirm identities about e^z for complex z, and by being able to factor any polynomial. For example, the double (and triple, half, etc.) angle formulas for sin and cos are easy to confirm using Euler's formula. Following easily from Euler's formula are the following important identities.

3 • 3. Result

For every real number θ and x,

1. $\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_x e^{ix} = i \cdot e^{ix}$;
2. $\cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) = \operatorname{Re}(e^{i\theta})$;
3. $\sin\theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta}) = \operatorname{Im}(e^{i\theta})$;

[i] "Euler" is pronounced as /ˈɔɪlər/, using the notation from the International Phonetic Alphabet.

This means that we can easily turn powers of $\cos\theta$ and $\sin\theta$ into powers and sums of $e^{i\theta}$ and $e^{-i\theta}$, which can be far more easy to integrate, as we'll see later. But beyond integration and e, the use of $\mathbb{C}$ is its ability to factor polynomials: *any* polynomial can be factored into factors of the form $(x-c)$ with c complex.

In $\mathbb{R}$, the polynomial x^2+4 cannot be factored. In $\mathbb{C}$, however, it can be rewritten as $x^2-(2i)^2$, which we know can be factored into $(x+2i)(x-2i)$. The fact that any polynomial can be factored into terms of the form $(x-c)$ is what makes the complex numbers so nice to work with in the context of algebra. As another example, we can find cube roots of 1 (that aren't just 1). Using cartesian coordinates, this is difficult:

$$z^3=1 \longleftrightarrow z^3-1=0 \longleftrightarrow (z-1)(1+z+z^2)=0.$$

Factoring $1+z+z^2$ isn't obvious. Using the polar idea is much easier:

$$z^3=1 \longleftrightarrow (re^{i\theta})^3=1 \longleftrightarrow r^3e^{i3\theta}=1 \longleftrightarrow r^3=1 \text{ and } 3\theta=2\pi k \text{ for some } k \text{ in } \mathbb{Z}.$$

Hence the solutions are $z=e^{2\pi ik/3}$ for k in $\mathbb{Z}$, which yields three solutions in total.

Section 4. *Worked-Out Examples*

First, let's calculate some integrals with traditional calculus methods.

4 • 1. Example

Calculate $\int e^{3x}x\,\mathrm{d}x$.

Solution .:.

Here our choice of f' and g aren't super clear, right? We need to fill out a table:

$$\begin{array}{ll} f'=e^{3x} & f=\frac{1}{3}e^{3x} \\ g'=1 & g=x \end{array}$$

Using this, can calculate by integration by parts (2 B • 2)

$$\int e^{3x}x\,\mathrm{d}x=\int f'g\,\mathrm{d}x=fg-\int fg'\,\mathrm{d}x$$
$$=\frac{1}{3}xe^{3x}-\int\frac{1}{3}e^{3x}\,\mathrm{d}x=\frac{1}{3}xe^{3x}-\frac{1}{9}e^{3x}.$$

4 • 2. Example

Calculate $\int x^2\sin x\,\mathrm{d}x$.

Solution .:.

Here our choice of f' and g are a little less clear, in that it doesn't seem so immediate what to choose. We again need to fill out a table:

$$\begin{array}{ll} f'=\sin x & f=-\cos x \\ g'=2x & g=x^2 \end{array}$$

Using this, can calculate by integration by parts (2 B • 2)

$$\int x^2\sin x\,\mathrm{d}x=\int f'g\,\mathrm{d}x=fg-\int fg'\,\mathrm{d}x$$
$$=(-\cos x)x^2-\int -2x\cos x\,\mathrm{d}x.$$

At this point, it's not entirely clear how to integrate $\int x\cos x\,\mathrm{d}x$. So we will need to use integration by parts (2 B • 2) again: first we fill out a table.

$$\begin{array}{ll} u'=\cos x & u=\sin x \\ v'=1 & v=x \end{array}$$

And so we can calculate $\int x \cos x \, \mathrm{d}x$ (and from there $\int -2x \cos x \, \mathrm{d}x$ to yield $\int x^2 \sin x \, \mathrm{d}x$) through

$$\begin{aligned}\int x \cos x \, \mathrm{d}x &= \int u'v \, \mathrm{d}x = uv - \int uv' \, \mathrm{d}x \\ &= x \sin x - \int \sin x = x \sin x + \cos x.\end{aligned}$$

Using this in the above calculation, we get

$$\begin{aligned}\int x^2 \sin x \, \mathrm{d}x &= (-\cos x)x^2 - \int -2x \cos x \, \mathrm{d}x \\ &= -x^2 \cos x + 2(x \sin x + \cos x) \\ &= (2 - x^2) \cos x + 2x \sin x.\end{aligned}$$

Using complex numbers and functions is very natural. Integrals that traditionally might have required remembering a trigonometric identity and using it cleverly can be made much more straight forward through the use of complex functions and Euler's formula (3 • 2). For example, consider the integral of $\cos^2 x$. Normally, you would need to use the trigonometric identity $\sin^2 x = 1 - \cos^2 x$. But if this were forgotten, a reasonable approach would be the following.

4 • 3. Example

Calculate $\int \cos^2 x \, \mathrm{d}x$.

Solution .:.

Here our choice of f' and g is pretty clear: both should be $\cos x$. But then we need to fill out a table:

$$\begin{aligned} f' &= \cos x & f &= \sin x \\ g' &= -\sin x & g &= \cos x \end{aligned}$$

Hence we get

$$\begin{aligned}\int \cos^2 x \, \mathrm{d}x &= \int f'g \, \mathrm{d}x = fg - \int fg' \, \mathrm{d}x \\ &= \sin x \cos x - \int \sin x(-\sin x) \, \mathrm{d}x \\ &= \sin x \cos x + \int \sin^2 x \, \mathrm{d}x.\end{aligned}$$

Now we want to find the integral of $\int \sin^2 x \, \mathrm{d}x$. So we again must do integration by parts, and fill out a table:

$$\begin{aligned} h' &= \sin x & h &= -\cos x \\ k' &= \cos x & k &= \sin x \end{aligned}$$

Hence we get

$$\begin{aligned}\int \sin^2 x \, \mathrm{d}x &= \int h'k \, \mathrm{d}x = hk - \int hk' \, \mathrm{d}x \\ &= -\cos x \sin x - \int (-\cos x) \cos x \, \mathrm{d}x \\ &= -\sin x \cos x + \int \cos^2 x \, \mathrm{d}x.\end{aligned}$$

Plugging this back into our formula for $\int \cos^2 x \, \mathrm{d}x$ gives us nothing new:

$$\int \cos^2 x \, \mathrm{d}x = \sin x \cos x - \sin x \cos x + \int \cos^2 x \, \mathrm{d}x = \int \cos^2 x \, \mathrm{d}x.$$

We can see that this approach fails. But where this method failed, using Euler's formula (3 • 2) works.

4 • 4. Example

Calculate $\int \cos^2 x \, \mathrm{d}x$.

Solution .:.

Recall that $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$. Hence when we square, we get the sum of easily integrable functions:

$$\begin{aligned}\int \cos^2 x \, dx &= \int \frac{1}{4}(e^{ix} + e^{-ix})^2 \, dx \\ &= \int \frac{1}{4}(e^{2ix} + 2 + e^{-2ix}) \, dx \\ &= \frac{1}{4}\left(\frac{1}{2i}e^{2ix} + 2x + \frac{-1}{2i}e^{-2ix}\right) + c\end{aligned}$$

In simplifying this, we can either use Euler's formula (3 • 2), or recognize $\sin 2x$ inside there. Either way gives the same result:

$$\int \cos^2 x \, dx = \frac{1}{4}(\sin(2x) + 2x) + c = \frac{1}{4}\sin(2x) + \frac{1}{2}x + c.$$

Of course, if one did remember the trigonometric identity, this way seems a little more involved. But a simple change from $\cos^2$ to $\cos^3$, or higher, makes the problem even more difficult when going the trigonometric route. Even if computationally tedious, using Euler's formula (3 • 2) is much more straightforward. We can consider another example comparing the two methods.

4 • 5. Example

Calculate $\int e^x \sin x \, dx$.

Solution .:.

Unlike the previous example, we don't need to remember many trigonometric identities to use the traditional method. We do, however, need to have some creativity and intuition about what to do. Using integration by parts, we will fill out the table

$$\begin{aligned}f' &= \sin x & f &= -\cos x \\ g' &= e^x & g &= e^x\end{aligned}$$

Hence we have the equality

$$\int e^x \sin x \, dx = \int f'g \, dx = fg - \int fg' \, dx = -e^x \cos x + \int e^x \cos x \, dx.$$

Now we need to calculate $\int e^x \cos x \, dx$. Again, the only real option available to us is integration by parts. Now we must again make some decisions on how to proceed, the second place intuition is required.

$$\begin{aligned}u' &= \cos x & u &= \sin x \\ v' &= e^x & v &= e^x\end{aligned}$$

Hence we have the equality

$$\int e^x \cos x \, dx = \int u'v \, dx = uv - \int uv' \, dx = e^x \sin x - \int e^x \sin x \, dx.$$

At this point, it may appear that we're stuck: in trying to calculate $\int e^x \sin x \, dx$, it seems like we need to have already calculated it. In the third occurance of intuition, we can realize the equality our first use of integration by parts gave us:

$$\begin{aligned}\int e^x \sin x \, dx &= -e^x \cos x + \int e^x \cos x \, dx \\ &= -e^x \cos x + e^x \sin x - \int e^x \sin x \, dx \\ \therefore 2\int e^x \sin x \, dx &= e^x \cos x + e^x \sin x + c \\ \int e^x \sin x \, dx &= \frac{1}{2}e^x \cos x + \frac{1}{2}e^x \sin x + c.\end{aligned}$$

The $+c$ comes from the fact that we're using indefinite integrals: the $\int e^x \sin x \, dx$ on the left may differ from $\int e^x \sin x \, dx$ on the right by a constant. But the important thing is that we found our answer through a long process.

A method which is much more mechanical, requiring fewer choices, and less intuition, is the following method:

4 • 6. Example

Calculate $\int e^x \sin x \, dx$.

Solution .:.

We can use either the identity of

$$\sin x = \operatorname{Im}(e^{ix}), \quad \text{or} \quad \sin x = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right).$$

Either will work, but to showcase a different method, we'll consider the identity $\sin x = \operatorname{Im}(e^{ix})$, as witnessed by Euler's formula (3 • 2). The is the only part of the calculation which requires some intuition. But what it means is that

$$\int e^x \sin x \, dx = \int e^x \operatorname{Im}(e^{ix}) \, dx = \int \operatorname{Im}(e^x e^{ix}) \, dx = \operatorname{Im}\left(\int e^x e^{ix} \, dx\right)$$
$$= \operatorname{Im}\left(\int e^{x(1+i)} \, dx\right) = \operatorname{Im}\left(\frac{1}{1+i} e^{x(1+i)}\right) + c.$$

So the work of calculating the integral is already done. The only work left is just expanding definitions and so forth to find the imaginary part. This means that we need to put the inside part $\frac{1}{1+i} e^{x(1+i)}$ into that cartesian form $a + bi$. Through the usual methods of expanding, using Euler's formula (3 • 2), and getting square roots out of the denominator,

$$\begin{aligned}
\frac{1}{1+i} e^{x(1+i)} &= \frac{1}{1+i} \cdot \frac{1-i}{1-i} e^x e^{ix} \\
&= \frac{1-i}{1-i^2} e^x e^{ix} \\
&= e^x \frac{1-i}{2} (\cos x + i \sin x) \\
&= \frac{1}{2} e^x (\cos x + i \sin x - i \cos x + \sin x) \\
&= \frac{1}{2} e^x (\cos x + \sin x) + i \frac{1}{2} e^x (\sin x - \cos x).
\end{aligned}$$

This is of the form $a + bi$, and so its imaginary part is $\frac{1}{2} e^x (\sin x - \cos x)$, which means

$$\int e^x \sin x \, dx = \frac{1}{2} e^x (\sin x - \cos x) + c.$$

So unlike the traditional method, there was only one creative decision: to use either $\operatorname{Im}(e^{ix})$ or Euler's formula (3 • 2) at the start. Everything else was just simplifying and using one memorized integral: $\int e^{ax} = e^{ax}/a + c$.

Now for an example of the use of complex numbers in factoring.

4 • 7. Example

Factor $x^4 + 3$ (the professor wants you to realize that things don't always come out as nice integers).

Solution .:.

Note that the factors $(x - c)$ of $x^4 + 3$ have $c^4 + 3 = 0$. So we're looking for solutions to $z^4 + 3 = 0$, i.e. $z^4 = -3$, for z in $\mathbb{C}$. In polar form, we're trying to find $r \geq 0$ and θ where

$$z^4 = (re^{i\theta})^4 = r^4 e^{4i\theta} = -3.$$

Writing $-3 = 3 \cdot e^{i\pi}$, we can equate the magnitudes and angles (modulo 2π). So we must have

$$|z^4| = r^4 = 3 = |-3|, \text{ and}$$
$$4\theta = \pi + 2\pi k, \text{ for some } k \text{ in } \mathbb{Z}.$$

Now the only real numbers with $r^4 = 3$ are $+\sqrt[4]{3}$ and $-\sqrt[4]{3}$. As r must be positive, it must be $\sqrt[4]{3}$. Hence our solutions are $\sqrt[4]{3} e^{i(1+2k)\pi/4}$ for k as 0, 1, 2, and 3.

Returning to the original question, the factorization is then

$$x^4 + 3 = \left(x - \sqrt[4]{3}e^{i\pi/4}\right)\left(x - \sqrt[4]{3}e^{i3\pi/4}\right)\left(x - \sqrt[4]{3}e^{i5\pi/4}\right)\left(x - \sqrt[4]{3}e^{i7\pi/4}\right).$$

4 • 8. Example

Use Euler's formula to write e^{2-3i} in the form $a + bi$ for a and b real numbers.

Solution .:.

$$e^{2-3i} = e^2e^{-3i} = e^2(\cos(-3) + i\sin(-3)) = e^2\cos 3 - ie^2\sin 3.$$

4 • 9. Example

Use Euler's formula to write 2^{1-i} in the form $a + bi$ for a and b real numbers.

Solution .:.

$$2^{1-i} = 2 \cdot 2^{-i} = 2e^{-i\ln(2)} = 2(\cos(-\ln 2) + i\sin(-\ln 2)) = 2\cos(\ln 2) - i \cdot 2\sin(\ln 2).$$

Chapter II. First-Order, Differential Equations

Section 5. Basics

Unknowingly, you've come across differential equations before in calculus. One of the simplest differential equations would be $y' = 0$, which implies that y is just a constant. Similarly, if $y' = a$ for some constant a, we get that $y(t) = at + c$ for some constant c.[i] Thus the solution to $y' = a$ is not unique: for different c, we get different $y = at + c$ satisfying the differential equation. In general, we've seen the differential equation $y' = f(t)$, in which case the general solution is the (indefinite) integral of f.

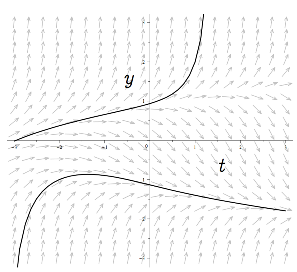
To generalize this, we will consider what happens when y itself is related to its derivative. This means differential equations of the form $y' = f(t, y)$ for some function f of t and y. Note that there might not be *any* solutions to an arbitrary differential equation. For example, $y' = y' + 1$ has no solution for y. Equations are statements about functions, and so are sometimes true and sometimes false. Our goal is to find out when they are true. This can be done by examining the equations themselves. Again, let's consider equations of the form $y' = f(t, y)$ for some function f. Using this, we can see what y' would need to be if a solution y had a value y_0 at t_0. Plotting this gives a *direction field*.

§ 5 A. Direction fields

Direction fields are primarily a visual aid in that they give a general idea how any given solution to a differential equation behaves. You can think of these fields as acting like currents of a river: if you plop in at some initial value, the path the direction field carries you is the solution[ii] with that initial value. These fields are easily displayed when we have y' equal to an expression involving just y and t:

$$y' = f(t, y),$$

for some function f. The idea is that each point (a, b) in the ty-plane then has a corresponding direction it goes according to the above equation. In practice, this means drawing a small line with slope $f(a, b)$ at (a, b).



5 A • 1. Figure: Direction field of $y' = y^2 + \frac{1}{2}y - \frac{1}{2}t - 1$

Again, this is mostly a visual aid, allowing you to see at a glance whether solutions will tend towards a certain value or not just by starting from a variety of initial values, and following the lines of the diagram. For example, consider the differential equation $y' = y^2 + \frac{1}{2}y - \frac{1}{2}t - 1$. If we consider the values of y' when y and t are between -3 and 3,

[i] In particular, this c is just $y(0) = a \cdot 0 + c = c$.

[ii] Later results tell us that there is in fact only one such solution (when f is continuously differentiable with respect to y) with that initial value, so the use of "the" here is appropriate.

we can get a variety of slopes. These slopes are then used as the slopes of arrows, giving the direction field in Figure 5 A • 1. Two solutions are also plotted to show what the solutions look like.

There are many different solutions to this differential equation, like most, since any given solution might have different *initial values*. For example, $y' = 1$ has solutions t, but also $t + 2$, and $t + c$ for any c. Remembering constants of integration will turn out to be very important here. Whereas in previous calculus courses such a constant wasn't very important, here it can be crucial.

§ 5 B. Descriptions of differential equations

Sometimes it's easy to solve certain kinds of differential equations. As such, it's useful to have names to reference such equations. Additionally, working with a smaller subset of differential equations can be easier than dealing with them in general. The same holds for us. For our purposes, we have the following concepts and descriptions.

5 B • 1. Definition

A differential equation

- is *ordinary* iff it doesn't involve partial derivatives.
- is *partial* iff it involves partial derivatives.
- is *nth order* iff it involves the nth derivative (without any higher derivatives).

The first kinds of differential equations we will be looking at will be *separable*, and *linear*.

5 B • 2. Definition

A differential equations is *separable* iff it is of the form (or at least can be put in the form)

$$y' = f(t)g(y),$$

for some functions f and g.

A general process for solving separable equations is then to divide by $g(y)$, yielding $y'/g(y) = f(t)$. Integrating then gives an equations involving just y and t, with no derivatives.

5 B • 3. Definition

A differential equation is *linear* iff it is of the form (or at least can be put in the form)

$$a_0(t)y + a_1(t)\frac{\mathrm{d}y}{\mathrm{d}x} + \cdots + a_n(t)\frac{\mathrm{d}^n y}{\mathrm{d}x^n} = g(t),$$

for some functions $g, a_0, \cdots, a_n$.

One simple test to see whether an equation is linear to see whether there are any terms like $y' \cdot y$ or $y'''' \cdot y''$. If there are—and you can't just divide by or subtract them to remove them from both sides—then the equation is not linear. Over time, we will look at more types of differential equations.

§ 5 C. Initial value problems

As above, there can be many solutions to a differential equation. Specifying an initial value just means specifying the solution. Usually, you can end up with a general form for the solution involving some unknowns like a $+c$ from an integral:

$$y(t) = f(t, c),$$

for some function f. If this is the general solution, specifying the initial value $y(0) = y_0$ requires

$$y_0 = f(0, c),$$

which often is enough to uniquely identify c. Note that 0 isn't particularly special, as we could just as easily specify the value at 1, $y(1) = y_1$, and so require c to satisfy

$$y_1 = f(1, c).$$

Either way, solving for c—and plugging this back into the general form—gives us an explicit way to calculate y for any value of t.

§ 5 D. Two simple initial value problems

One of the simplest examples of interaction between y' and y is just equality.

$$y' = y.$$

This equation is both separable, and linear, and so the solutions of this can be found just by dividing by y and then integrating. If $y \neq 0$ somewhere, then

$$y' = y \longleftrightarrow \frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}t} = 1 \longleftrightarrow \int \frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}t}\,\mathrm{d}t = 1 \longleftrightarrow \int \frac{1}{y}\,\mathrm{d}y = 1 \longleftrightarrow \ln|y| + C_1 = t + C_2$$
$$\longleftrightarrow \ln|y| = t + C_3 \longleftrightarrow |y| = e^{t+C_3} = e^{C_3}e^t.$$

By choosing a new constant[iii]

$$C = \begin{cases} e^{C_3} & \text{if } y > 0 \text{ somewhere} \\ -e^{C_3} & \text{if } y < 0 \text{ somewhere} \\ 0 & \text{if } y = 0 \text{ somewhere.} \end{cases}$$

we can get rid of the absolute value around y to get the equality

$$y = Ce^t,$$

for some constant C. In fact, note that $y(0) = Ce^0 = C$, so that we can actually say $y = y(0)e^t$. Such a derivation is common with simple equations like this, and it is one of the reasons we will encounter e so much.

Let's consider a slightly more complicated relationship between y' and y:

$$y' = ay - b.$$

For various *numbers* a and b. Later we'll encounter more complicated expressions where $y' = f(t)y - g(t)$ for functions f and g. For now, this simpler equation can be solved in the same way each time, in a similar way as before. In special cases, this complicated expression for y can be simplified quite a bit, e.g. when t_0 or w is 0.

5 D • 1. Result

Let $a \neq 0$. Let y be a function satisfying $y' = ay - b$, with $y(t_0) = w$, some specified value. Therefore for any t,

$$y(t) = \left(w - \frac{b}{a}\right)e^{a(t-t_0)} + \frac{b}{a}.$$

Proof .:.

Note that $y' = ay - b$ implies $y' = a(y - b/a)$, and hence

$$\frac{y'}{y - b/a} = a.$$

If we integrate both sides with respect to t, we get

$$\int \frac{y'(t)}{y(t) - b/a}\,\mathrm{d}t = at + c$$

for some constant c. Through u-substitution, we can evaluate the integral on the left. So for all t,

$$\ln|y(t) - b/a| = at + c.$$

We can then solve for $y(t)$ to get that $|y(t) - b/a| = e^c e^{at}$ and so $y(t) = Ce^{at} + b/a$. This C is just some constant—in particular, it's short-hand for $e^c \cdot \mathrm{sign}(w - b/a)$ given the c above. Calculating this C can be done with the initial value w:

$$w = y(t_0) = Ce^{at_0} + b/a \quad \longrightarrow \quad \frac{w - b/a}{e^{at_0}} = C.$$

[iii] This is a constant, since if $y > 0$ somewhere, then $y > 0$ everywhere, and similarly for $y < 0$. To see this, since y is continuous, otherwise y would need to go from being positive to negative, and so be 0 somewhere. But for y to be 0 somewhere requires y to be 0 everywhere by some later results about the uniqueness of solutions.

Hence, we can write

$$y(t) = \left(w - \frac{b}{a}\right) e^{a(t-t_0)} + \frac{b}{a},$$

which of course, presupposes $a \neq 0$. If $a = 0$, then the solution is trivial: $y' = b$ implies y is just a line with slope b through the point (t_0, w). ⊣

Problems involving these differential equations come in a few forms. The first is to go through the process above in a particular case. Another is to find a value of t where $y(t) = m$ for some value m. This can be done by finding the solution as above, and then solving for t in the equation $y(t) = m$.

§ 5 E. Existence and uniqueness theorems

The existence and uniqueness theorems are important for actual mathematics, but here we're focused on doing calculations rather than doing math. So for our purposes, these theorems merely confirm that we don't need to do extra work in finding solutions. Once we've found a solution that works, that's the *only* solution: there's no need to look any further to see if maybe we're missing something. Of course, the theorems don't apply in all cases, so we must be careful in general.

Of the several results the book[1] gives, the existence and uniqueness of a solution to a differential equation is satisfied so long as the given functions are continuous—or in some sense smooth—in a given area, which guarantees the existence and uniqueness for at least a part of that area.

5 E • 1. Result (The Existence and Uniqueness Theorem)

Let f and $\partial f/\partial y$ be continuous in some rectangle R. Therefore there is exactly one y such that $y(t_0) = y_0$ and

$$y' = f(y, t),$$

where y is defined on an (open) interval around t_0 within the bounds of R.

This means that we have existence and uniqueness locally, although perhaps not globally. As an analogy, the equation for (the graph of) the unit circle is $x^2 + y^2 = 1$. This does not define a function y, since there will be two ys for any $x \neq \pm 1$. But if we restrict our view to x in the interval $[-1, 1]$ and y in the interval $[0, 1]$, it does define a function: $y = \sqrt{1 - x^2}$ over $(-1, 1)$.

Section 6. Linear Equations

In the previous section, we worked with differential equations of the form

$$y' = ay + b,$$

for constants $a \neq 0$[iv] and b. This can be generalized slightly, where a and b are instead functions: $y' = f(t)y + g(t)$. In a slightly more usable form, we can write how this will usually be written:

$$y' + p(t)y = g(t)$$

for functions p and g. Unlike the simpler equation, this is not separable, so our method to solve the equation cannot be so simple as just (dividing and then) integrating both sides. Instead we have two methods of finding solutions.

§ 6 A. Integrating factors

The method of integrating factors is just finding a $\mu \neq 0$ that satisfies

$$(\mu(t)y)' = \mu(t)y' + \mu(t)p(t)y, \qquad \text{i.e. that satisfies} \qquad \mu'(t) = \mu(t)p(t).$$

[iv] if $a = 0$, the differential equation is just $y' = b$, which is easily solved: $y(t) = bt + y(0)$.

This method allows us to conclude the following.

6 A • 1. Result

Let p and g be continuous functions, and consider the differential equation $y' + p(t)y = g(t)$ with $y(t_0) = y_0$. Therefore

$$\mu(t) = e^{\int_{t_0}^{t} p(x)\,\mathrm{d}x}$$

is an integrating factor, and

$$y(t) = \frac{y_0}{\mu(t)} + \frac{1}{\mu(t)} \int_{t_0}^{t} \mu(x)g(x)\,\mathrm{d}x.$$

Proof .:.

The key observation is that $y' + p(t)y$ looks a bit like the product rule:

$$(\mu(t)y)' = \mu(t)y' + \mu'(t)y.$$

If we can find a μ where $\mu' = \mu p(t)$, then the equation becomes a simple matter of integrating both sides again, and then dividing by $\mu(t)$. To solve $\mu' = \mu p(t)$, note that then

$$\frac{\mu'}{\mu} = p(t).$$

Hence the integrals are equal: $\int_{t_0}^{t} \mu'/\mu\,\mathrm{d}t = \int_{t_0}^{t} p(t)\,\mathrm{d}t$. As a result, a u-substitution of $\mu(t)$ yields

$$\ln|\mu(t)| = \int_{t_0}^{t} p(t)\,\mathrm{d}t + c$$

for some constant c. Note that then

$$\mu(t) = e^{\int_{t_0}^{t} p(t)\,\mathrm{d}t} \cdot C$$

is then the solution, for some value of C. It's important to realize now that we were just trying to find *some* function μ that would satisfy the equation. So any choice of C, excluding 0, works. To make things simpler, we will choose $C = 1$, making $\mu(t_0) = C = 1$. As a result $\mu(t)$ is never 0, and so we can divide and multiply by it:

$$\begin{aligned} g(t) = y' + p(t)y &\longleftrightarrow \mu(t)g(t) = \mu(t)y' + \mu(t)p(t)y \longleftrightarrow \mu(t)g(t) = \mu(t)y' + \mu'(t)y \\ &\longleftrightarrow \mu(t)g(t) = (\mu(t)y)' \longleftrightarrow \int_{t_0}^{t} \mu(t)g(t)\,\mathrm{d}t = \mu(t)y(t) - \mu(t_0)y_0 \\ &\longleftrightarrow \int_{t_0}^{t} \mu(t)g(t)\,\mathrm{d}t = \mu(t)y(t) - y_0 \longleftrightarrow y_0 + \int_{t_0}^{t} \mu(t)g(t)\,\mathrm{d}t = \mu(t)y(t) \\ &\longleftrightarrow \frac{y_0}{\mu(t)} + \frac{1}{\mu(t)} \int_{t_0}^{t} \mu(t)g(t)\,\mathrm{d}t = y(t). \end{aligned}$$

Substituting the definition of μ then gives the desired conclusion. ⊣

As a result, the method of integrating factors allows one to solve all first-order, linear, ordinary differential equations.[v]

§ 6 B. First-order undetermined coefficients

Integrating factors can be a slow process, however, requiring us to integrate potentially very complicated functions. A slightly easier idea is contained in a kind of guess and check. This method is further elaborated on in Subsection 14 B. For now, consider the version of the differential equation where $g(t) = 0$, and $p(t) = a$ is some constant.

$$y' + ay = 0.$$

Note that any two solutions to $y' + ay = g(t)$ will have their difference satisfy $y' + ay = 0$. As a result, if $\phi(t)$ is the general solution to $y' + ay = 0$, and $Y(t)$ is *any* solution to $y' + ay = g(t)$, then we can say that the general solution to $y' + ay = g(t)$ will be $y(t) = \phi(t) + Y(t)$. Solving $y' + ay = 0$ is relatively easy, since it's separable: $y' = -ay$ and hence the general solution is $y(t) = Ce^{-at}$.

[v] so long as the functions involved are sufficiently nice, which they will be in almost all practical applications.

To find a particular solution $Y(t)$, we can do guess-and-check based on $g(t)$. This is more restrictive than integrating factors, but can sometimes be much easier to do in practice.

6 B • 1. Result (Undetermined Coefficients)

Consider the differential equation $ay' + by = g(t)$.

If $g(t) = a_n t^n + \cdots + a_1 t^1 + a_0$,	then guess	$Y(t) = A_n t^n + \cdots + A_1 t^1 + A_0$;
if $g(t) = \sin(at)$, or $g(t) = \cos(at)$,	then guess	$Y(t) = A\sin(at) + B\cos(at)$;
if $g(t) = e^{at}$,	then guess	$Y(t) = Ae^{at}$;

for various constants.

If the guess ends up being a solution to $ay' + by = 0$, then just multiply the guess by t. And if that also is a solution, keep multiplying by t until it isn't.

If g is a product/sum of these, then our guess should be a product/sum of the corresponding guesses. For example, if $g(t) = t\cos(2t) + 5e^t$, we should guess

$$Y(t) = (At + B)(C\cos(2t) + D\sin(2t)) + Ee^t,$$

for undetermined coefficients A, B, C, D, and E. Note we need to use different constants for each guess involved.

Once we have our guess, it's a matter of figuring out what constants work, which means plugging in $Y(t)$ into $ay' + by = g(t)$, and solving for the coefficients of Y.

Because this is just one equation with many unknowns, it's not obvious how we can solve for many different coefficients. To get around this, note that such an equation is supposed to hold for all t, and thus we can plug in various values of t to get as many equations as we need. Alternatively—and more easily—we can equate coefficients, and get equations that way.

Section 7. Separable and Autonomous Equations

§ 7 A. Separable equations

There's another way to generalize the differential equations worked with in Section 5:

$$y' = h(y) \cdot g(t).$$

In the case of $y' = ay + b$, $g(t)$ is just a constant. By dividing by $h(y)$, when this is non-0, a more workable form for the equation is

$$f(y)y' = g(t).$$

Such equations are called *separable*, since we have, in some sense, separated the variables y and t. Solving the differential equation is then just done by integration: using a u-substitution of y: $\int f(y)\,dy = \int g(t)\,dt$. In practice, we have the following result.

7 A • 1. Result

Let f and g be continuous functions such that

1. There is a function y satisfying $f(y)y' = g(t)$; and
2. $y(t_0) = y_0$.

Therefore $F(y(t)) - F(y_0) = \displaystyle\int_{t_0}^{t} g(t)\,dt$.

Proof .:.

By integrating both sides, we get through a u-substitution of y,

$$f(y)y' = g(t) \longleftrightarrow \int_{t_0}^{t} f(y)y' \, dt = \int_{t_0}^{t} g(t) \, dt \longleftrightarrow \int_{y_0}^{y(t)} f(y) \, dy = \int_{t_0}^{t} g(t) \, dt$$

$$\longleftrightarrow F(y(t)) - F(y_0) = \int_{t_0}^{t} g(t) \, dt. \qquad \dashv$$

Because the method is so simple, separable equations are useful to have. Unfortunately, the integration of the functions can be difficult. Similarly, taking the inverse of F to get a unique y might also be difficult to achieve, or possibly impossible. Often it can be useful to think of y not as a function, but as another variable in relation to t, in particular one of potentially many y satisfying

$$F(y) = F(y_0) + \int_{t_0}^{t} g(t) \, dt.$$

The resulting graph is then a path through the plane, but it might not be the graph of a function—i.e. it fails the vertical line test.

§ 7 B. Autonomous equations

To make things a little simpler, we can consider differential equations where there is no interaction from t.

7 B • 1. Definition

A differential equation is *autonomous* iff it is of the form $y' = f(y)$ for some function f.

This means autonomous equations are separable, and so we can apply the methods there to solve the differential equations. Much of the book[1], however, is mostly focused on the stability of various constant solutions. In essence, if we find a constant c such that $f(c) = 0$, then the constant function $y(t) = c$ satisfies the differential equation above: $y'(t) = 0$, and $f(y(t)) = f(c) = 0$ so that $y' = f(y)$. This doesn't necessarily hold in general, since in other cases y' might depend on t. But for *autonomous* equations, y' only depends on y.

7 B • 2. Definition

Consider the differential equation $y' = f(y)$ for some f. A *critical point* of f, or *equilibrium* is a constant solution $y = c$. The critical point $y = c$ is

1. *stable* iff solutions with initial values "near" c approach c as $t \to \infty$.
2. *unstable* iff solutions with initial values "near" c diverge from c as $t \to \infty$.
3. *semistable* iff solutions near c converge on one side, and diverge on the other.

Being "near" c is a pretty ambiguous statement, but in most cases it's clear what's meant. For instance, if we have the equation

$$y' = (y-1)^2(y-2)(y-3),$$

then we have three constant solutions: $y(t) = 1$, $y(t) = 2$, and $y(t) = 3$. Being "near" the solution $y(t) = 2$ in this case just means being somewhere between 1 and 2, or 2 and 3. But being "near" $y(t) = 1$ means just being anywhere from $-\infty$ to 1, or from 1 to 2.

Finding out whether a given constant solution is stable, unstable, or semistable is often done just by looking at whether y' is positive or negative near that value. Continuing the previous example,

$$y' = (y-1)^2(y-2)(y-3)$$

is positive for $y < 1$. Hence solutions near 1 converge to 1 when they are below it. Solutions between 1 and 2 yield y' as positive. Thus solutions near 1 diverge from 1 when they are above it. Since they converge from below, and diverge when above, $y = 1$ is a semistable equilibrium.

Section 8. *Exact Equations*

Exact equations are a kind of generalization of separable equations, but where separable meant something of the form

$$f(y)y' + g(t) = 0,$$

exact equations allow f and g to be functions of both y and t:

$$f(y,t)y' + g(y,t) = 0,$$

where $f(y,t)$ and $g(y,t)$ are partial derivatives of some other given function.

8 • 1. Definition

A differential equation is *exact* iff it can be put in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}F(y,t) = \frac{\partial F}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial F}{\partial t} = 0,$$

for some $F(y,t)$.

From the multi-variable chain rule, this just states that

$$\frac{\mathrm{d}}{\mathrm{d}t}F(y,t) = 0, \qquad \text{i.e.} \qquad F(y,t) = c,$$

for some constant c. Solving the differential equation is then the same as solving $F(y,t) = c$ for y.

Just like with separable equations, it might not be possible to get a unique y from this. Nevertheless, such a form can allow us to describe solutions. More important for our purposes is how to find such an F. Supposing that we're given nice enough functions, we can apply the following result which tells us when an equation is exact, and how to find the function in question. Note that we're denoting partial derivatives by subscripts here: $f_t = \partial f/\partial t$ for example.

8 • 2. Result

Let f and g be continuously differentiable functions. Therefore

$$f(y,t)y' + g(y,t) = 0 \text{ is exact} \qquad \text{iff} \qquad f_t = g_y.$$

Figuring out whether the equation is exact can be done fairly easily thanks to this pretty simple test: the equation is exact iff $f_y = g_t$, referring to partial derivatives again. Once we know that an equation is exact, finding such an F is a problem of integration. In essence, we have

$$F = \int f(y,t)\,\mathrm{d}t + c(y).$$

So now we regard the (no-longer) constant c from integration as a function of y. To find this $c(y)$, we integrate $g(y,t)$:

$$c(y) = \int g(y,t)\,\mathrm{d}t - \int f(y,t)\,\mathrm{d}t.$$

There's nothing special about f versus g here. We could just as easily do the reverse order, but the point is the equality

$$F(y,t) = \int f(y,t)\,\mathrm{d}t + c_1(y) = \int g(y,t)\,\mathrm{d}y + c_2(t),$$

for functions c_1 and c_2. This also makes intuitive sense since taking the partial derivatives makes these c_1 and c_2s disappear, leaving just $f(y,t)$ or $g(y,t)$. Note again that such an F is not unique, since $F+2$ or $F+15$ works just as well as a witness to the fact that $f + gy' = 0$ is an exact differential equation. But we're not interested in finding a unique F, since just one will suffice to solve the differential equation.[vi]

The difficulty with exact equations is that they are really rare. But once you know that the equation *is* exact, then the differential equation can be dealt with fairly easily.[vii] Now sometimes we can turn a non-exact differential equation into an exact differential equation through the method of integrating factors: multiplying by some non-0 $\mu(t)$—solving another differential equation to find such a μ—and then solving the new exact equation. The idea is similar to the integrating factors used for linear equations, except now more general.

[vi] and the difference of a constant won't matter, since this constant could be subtracted from the equation $F(y,t) = c$ to yield merely a different arbitrary constant on the right hand side.

[vii] at least conceptually

8 • 3. Result

Consider the non-exact differential equation

$$M(t, y)y' + N(t, y) = 0.$$

An integrating factor μ that turns this into an exact differential equation is one where

$$(\mu M)y' + (\mu N) = 0,$$

and which satisfies the partial differential equation

$$\mu_t M + \mu M_t - \mu_y N - \mu N_y = 0.$$

This result allows us to conclude that if μ is a function of just t, then

$$\mu_t = \mu' = \mu \frac{N_y - M_t}{M},$$

allowing us to solve the differential equation to find an integrating factor μ. Similarly , if μ is a function of just y, then

$$\mu_y = \mu' = \mu \frac{M_t - N_y}{N}.$$

Often these can be useful for confirming that we can find such a μ by instead seeing whether something like

$$\frac{M_t - N_y}{N}$$

is a function of just t or just y. But most of the time, finding such an integrating factor is not at all easy, and the above ideas won't always work.

Section 9. *Exercises*

II • Ex1. Exercise: Consider the differential equation $y' + 2y = e^{-2t}$ with initial condition $y(0) = 2$. Solve for $y(t)$ using undetermined coefficients.

II • Ex2. Exercise: $S(0) = -8000$, and $S' = .1S + k$ for some constant k:
8000 dollars is borrowed at an annual interest rate of 10%, compounded continuously with continuous payments at a rate of k.

1. Calculate k if $S(3) = 0$.
2. Calculate the amount paid during the 3 year period.

II • Ex3. Exercise: $S(0) = -150,000$, and $S' = \frac{.06}{12}S + (800 + 10t)$:
150,000 dollars is borrowed at an annual interest rate of 6%, compounded continuously with continuous payments at a rate of $800 + 10t$ per month.

a. When is $S(t) = 0$?
b. How large of a loan could be paid off in exactly 20 years given the payment of $800 + 10t$ per month?

II • Ex4. Exercise: Suppose that $u' = -\alpha u^4$ with $u(0) = 2000$, and $\alpha = 2 \cdot 10^{-12}$.

1. Solve the differential equation.
2. Find the time τ where $u(\tau) = 600$.

II • Ex5. Exercise: Consider the differential equation $(t - 3)y' + y \ln t = 2t$ where $y(1) = 2$. Determine an interval where the solution this initial value problem is certain to exist.

II • Ex6. Exercise: Consider the differential equation $y' = (t^2 + y^2)^{3/2}$.
State where in the ty-plane the hypotheses of The Existence and Uniqueness Theorem (5 E • 1) are satisfied.

II • Ex7. Exercise:

1. Verify that both $y_1(t) = 1 - t$ and $y_2(t) = -t^2/4$ are solutions of the intial value problem
$$y' = \frac{-t + \sqrt{t^2 + 4y}}{2}, \qquad y(2) = -1.$$
Where are these solutions valid?
2. Explain why the existence of two solutions of the given problem does not contradict the uniqueness part of The Existence and Uniqueness Theorem (5 E • 1).
3. Show that $y = ct + c^2$, where c is an arbitrary constant, satisfies the differential equation in part a for $t \geq -2c$. If $c = -1$, the initial condition is also satisfied, and the solution $y = y_1(t)$ is obtained. Show that there is no choice of c that gives the second solution $y_2(t)$.

II • Ex8. Exercise: Suppose $y' = -\beta y$, and $x' = -\alpha x y$. x is the proportion of susceptibles, y is the proportion of carriers.

1. Solve for $y(t)$ given that $y(0) = y_0$.
2. Solve for $x(t)$ given that $x(0) = x_0$ using $y(t)$ from part a.
3. Find the proportion of the population that escapes the epidemic by finding the limiting value of x as $t \to \infty$.

II • Ex9. Exercise: Consider $y' = e^{2x} + y - 1$. Find an integrating factor and solve the given equation.

II • Ex10. Exercise: Solve the equation
$$\frac{\mathrm{d}y}{\mathrm{d}t} + (1 + 2t)y = 3t, \qquad y(0) = 1.$$

II • Ex11. Exercise:

(a) Find the critical points of the equation
$$\frac{\mathrm{d}y}{\mathrm{d}t} = y \cdot (2y - 1) \cdot (y - 2).$$
(b) Find the type of each critical point (stable, unstable, semi-stable).

(c) If $y(0) = 1$ what is $\lim_{t\to\infty} y(t)$?
If $y(0) = 1/4$ what is $\lim_{t\to\infty} y(t)$?
If $y(0) = 3/2$ what is $\lim_{t\to\infty} y(t)$?

II • Ex12. Exercise:

1. Find the critical points of the equation $y' = y(1 + y^2 - 5y)$.
2. Find the type of each critical point.
3. If $y(0) = 1$, what is $\lim_{t\to\infty} y(t)$?

II • Ex13. Exercise: Solve the equation
$$y' = \left(1 + \frac{1}{t}\right) y + \begin{cases} 1 & \text{if } 1 \leq t \leq 2 \\ 2 & \text{if } 2 < t, \end{cases}$$
with initial condition $y(1) = 1$.

Chapter III. Discrete Problems

Section 10. *First-order Difference Equations*

Let's take a break from looking at continuous functions to looking at discrete problems. In particular, sequences $\langle y_0, y_1, y_2, \cdots \rangle$ defined by recursion, meaning that we get the next value from the previous ones:

$$y_{n+1} = f(n, y_n),$$

for some function f. This is supposed to be analogous to the differential equation $y' = f(t, y)$. Now clearly given an initial value y_0, the above equation uniquely defines a function: we just compute $f(0, y_0)$ to get y_1, and then we compute $f(1, y_1)$ to get y_2, $f(2, y_2)$ for y_3, and so on. But this process is both lengthy, and uninformative about the general trends of solutions.

The interest with these equations $y_{n+1} = f(n, y_n)$ is giving a *closed form* solution, meaning an expression that calculates y_n at any n (given the initial value y_0). This would allow us to see what happens to y_n as $n \to \infty$, for instance. If we can still find a closed form solution for arbitrary y_0, we can also see what happens if we vary the initial value y_0. In general, however, finding such closed form solutions is either impossible, or difficult at the very least. Although we can calculate any given value just by calculating f a bunch of times, it's hard to figure out how to calculate just based on n without knowing previous values.

There is no sure-fire way to find a closed form solution. Generally, the best method is to calcuate the first several values in terms of y_0, and then try to generalize. Of course, you then want to confirm that your solution $g(n)$ is actually correct by plugging it into f, and making sure $f(n, g(n))$ is just $g(n + 1)$ for all n.

Now there are more analogous properties of difference equations and differential equations.[i] For example, like with autonomous differential equations, we have a notion of *stability* of constant solutions (solutions where $y_n = c$ for all n, i.e. the sequence $\langle y_0, y_1, y_2, \cdots \rangle = \langle c, c, c, \cdots \rangle$).

10 • 1. Definition

Consider the first-order difference equation $y_{n+1} = f(y_n)$ for some f. The constant solution or *equilibirum* c is

1. *stable* iff solutions with y_0 "near" c approach c as $t \to \infty$.
2. *unstable* iff solutions with y_0 "near" c diverge from c as $t \to \infty$.

However, being "near" c is vague, even more-so than in autonomous differential equations (equations of the form $y' = f(y)$ for some f). In many cases it may be unclear where a solution is stable.

Rather than attempt to find an exact solution, it can be beneficial to instead *approximate* a solution. Given that the term "approximate" is a bit loose, there are a variety of ways to carry out such an approximation. Moreover, these different ways have their strengths in weaknesses: some are computationally easy but less accurate, and some are just the opposite. Some work better with certain kinds of difference equations, and some are more general. The methods covered here do not comprise a comprehensive list. They are really just some basic ideas that can get you started on thinking about approximating solutions.

[i] Well, really it's the definitions that are analogous.

Section 11. *Euler's Method*

§ 11 A. The method

Euler's method is a pretty intuitive way of approximating solutions. The idea is a step-by-step process where you look at what direction you're supposed to go, and then go a few steps in that direction, and then re-evaluate what direction you're supposed to go, and so on. The direction is determined by y', but the steps you take are arbitrary. But intuitively, if you keep correcting your direction really frequently—taking only small steps—you'll be more accurate in the end. This unfortunately means that if you want an accurate approximation, you'll need to take a lot of time to do the process over and over many times.

It's usually the case that the step size is the same every time: some value h. If you step by h every time, then you're approximating the value of the function at $y_0 = y(t_0)$, $y_1 = y(t_0+h)$, $y_2 = y(t_0+2h)$, and so on: $y_n = y(t_0+nh)$. This creates a first-order difference equation.

11 A • 1. Result

Suppose $y' = f(t, y)$ and $y(t_0) = y_0$. Thus Euler's method gives the first-order difference equation

$$y_{n+1} = y_0 + h \cdot f(t_0 + nh, y_n),$$

which is supposed to approximate the solution to the differential equation at $t_0 + nh$:

$$y(t_0 + nh) \approx y_n.$$

The resulting table of values of (t_1, y_1), (t_2, y_2), etc. yields an approximation of the solution to $y' = f(y, t)$ with initial value $y(t_0) = y_0$. Again, this process is long if done by hand, so such problems almost *need* a calculator if you don't want to waste your time.

§ 11 B. Errors in Euler's method

Recall that Euler's method is a step-by-step process where you look at where you're supposed to go, then go a few steps in that direction, and then re-evaluate what direction you're supposed to go, over and over. Now if we can actually solve the initial value problem $y' = f(t, y)$ with $y(t_0) = y_0$ for y, then we can look at the difference between the approximated values, and the *actual* values: the error in our approximations. Even better, if we can get a closed form solution to the difference equation in Result 11 A • 1, then we can look at how our error develops over time, and how good our approximated solution becomes.

11 B • 1. Definition

Consider the differential equation $y' = f(t, y)$ with $y(t_0) = y_0$.
Let $\phi(t)$ be a solution this.
Let $\langle y_0, y_1, y_2, \cdots\rangle$ be a sequence approximating ϕ by Euler's method with step size h.
The *error* or *global truncation error* of this approximation is the sequence $\langle E_0, E_1, E_2, \cdots\rangle$ where

$$E_n := \phi(t_n) - y_n.$$

The *local truncation error* of the method is the sequence $\langle e_0, e_1, e_2, \cdots\rangle$ would be the errors of the next step in Euler's method if we correctly started with $\phi(t_{n-1})$:

$$e_n := \phi(t_n) - \big[\phi(t_{n-1}) + h \cdot f(t_n, \phi(t_{n-1}))\big], \text{ while}$$
$$E_n := \phi(t_n) - \big[y_{n-1} + h \cdot f(t_n, y_{n-1})\big].$$

Errors in E_n have accumulated, but the errors on e_n have not. Given that $y_n \approx \phi(t_n)$ for all n, we should have that $e_n \approx E_n$ (as the above would suggest), but this is often not the case, especially for large n. The book[1] has its own favorite way of calculating the local truncation error with the following.

11 B • 2. Definition

Consider the differential equation $y' = f(t, y)$ with $y(t_0) = y_0$.
Let $\phi(t)$ be a solution to this with continuous second derivative ϕ''.
Consider Euler's method with step size h: $t_n = t + nh$. The *local truncation error* of this method is the sequence $\langle e_0, e_1, e_2, \cdots \rangle$ where

$$e_n := \frac{1}{2}\phi''(\bar{t})h^2,$$

where $\bar{t}$ is a fixed, unknown (to us) value between t_n and t_{n+1}.

The value of $\bar{t}$ is one is given by a cut-off taylor polynomial. In particular, we approximate ϕ with its taylor series, stopping at ϕ''. This isn't exactly ϕ, however, so we allow the ϕ'' term to be evaluated somewhere between t_n and t_{n+1} to ensure the equality actually holds:

$$\phi(t_{n+1}) = \phi(t_n) + \phi'(t_n)h + \frac{1}{2}\phi''(\bar{t})h^2.$$

This can be done so long as ϕ'' is continuous between t_n and t_{n+1}.

§ 11 C. Improvements in Euler's method

The improved Euler method is to approximate based on two things: where you are, and where you would end up. This requires solving an equation to get your result, but the approximation will be better. In particular,

11 C • 1. Definition

Consider the differential equation $y' = f(t, y)$ for some function f and initial value $y(t_0) = y_0$.
Let $t_n = t_0 + nh$ for some fixed h.
The *improved Euler method* with step-size h is a sequence $\langle y_0, y_1, \cdots \rangle$ satisfying

$$y_{n+1} = y_n + h\frac{f_n + f(t_{n+1}, y_n + hf_n)}{2}, \qquad \text{where } f_n := f(t_n, y_n).$$

The sequence is supposed to approximate the actual solution ϕ at the points t_0, t_1, and so on:

$$\begin{aligned} y_0 &= \phi(t_0) \\ y_1 &\approx \phi(t_1) \\ y_2 &\approx \phi(t_2) \\ y_3 &\approx \phi(t_3) \\ &\vdots \end{aligned}$$

Note that h and t_0 determine t_n for every n, so we can't skip values in our calculations. If $t_0 = 0$ and $h = .025$, to approximate $\phi(25)$, we would have to calculate $t_1, t_2, \cdots$, up to t_{1000}.

In essence, the improved Euler method is similar to using trapezoids in Riemann integration as opposed to rectangles.

Section 12. *Exercises*

III • Ex1. Exercise: Consider the differential equation with initial condition $y(0) = 1$

$$y' = 2y - 1.$$

a. Approximate y at $t = .1, .2, .3$, and $.4$ using Euler's method with $h = .1$.
b. Approximate y at those same values with with $h = .05$.
c. Approximate y at those same values with $h = .025$.

d. Find the solution to the differential equation, and compute its values at .1, .2, .3, and .4, comparing with the approximations.

III • Ex2. Exercise: Solve the difference equation

$$y_{n+1} = -\frac{9}{10} y_n$$

in terms of y_0, and describe the behavior of the solution as $n \to \infty$.

III • Ex3. Exercise: Solve the difference equation

$$y_{n+1} = (-1)^{n+1} y_n$$

in terms of y_0, and describe the behavior of the solution as $n \to \infty$.

III • Ex4. Exercise: Solve the difference equation

$$y_{n+1} = \frac{1}{2} y_n + 6$$

in terms of y_0, and describe the behavior of the solution as $n \to \infty$.

III • Ex5. Exercise: Suppose $y_0 = -100{,}000$ with y_{n+1} for $n \geq 0$ calculated by

$$y_{n+1} = \left(1 + \frac{9}{100} \cdot \frac{1}{12}\right) y_n + p$$

for some constant p (here n is in months).

1. What p (monthly payment) has $y_{30 \cdot 12} = 0$?
2. What p has $y_{20 \cdot 12} = 0$?
3. What is the total amount paid in each case?

III • Ex6. Exercise: Consider the differential equation

$$y' = \frac{1}{2} - t + 2y$$

with $y(0) = 1$.

Find approximate values of the solution at $t = .5$, 1, 1.5, and 2 using Euler's method with $h = .025$.

III • Ex7. Exercise: Obtain a formula for the local truncation error for Euler's method in terms of t and the exact solution $y = \phi(t)$ for the differential equation

$$y' = 5t - 3\sqrt{y}, \qquad y(0) = 2.$$

Chapter IV. Second-Order, Linear, Differential Equations

First order equations had many different types that we studied, but here we will be focused on two types: homogeneous, and non-homogeneous equations, both of which are kinds of linear equations. The primary kinds of equations we will be working with will be of the form

$$y'' + p(t)y' + q(t)y = g(t)$$

for functions p, q, and g. First, we need to consider what happens when g is 0.

Section 13. *Homogeneous Equations*

There is little connection with "homogenous" used in differential equations, and the term used in the rest of mathematics. For our purposes, we have the following definition.

13 • 1. Definition

A second-order differential equation is *homogeneous* iff it is of the form

$$P(t)y'' + Q(t)y' + R(t)y = 0,$$

for some functions P, Q, and R.

In other words, it is homogeneous iff it is linear with no constant (with respect to y) term. The solutions of these equations will be of the form $C\phi_1 + D\phi_2$ for solutions ϕ_1 and ϕ_2 where C and D are determined using the initial values of any given initial value problem: y_0 and y'_0. The expression $C\phi_1 + D\phi_2$ is a *linear combination*, a phrase that may be thrown around a lot at this point. We have the following definition

13 • 2. Definition

Let $\phi_1, \phi_2, \cdots, \phi_n$ be some things (functions, vectors, etc.). A *linear combination* of these is anything of the form

$$c_1\phi_2 + c_2\phi_2 + \cdots + c_n\phi_n$$

for constants c_1, c_2, $\cdots$, and c_n, the *coefficients*.

Finding solutions can be more difficult, but the idea is that we only need to find two solutions to get the general solution. The key thing, however, is that these solutions must be different enough. What characterizes this is having a non-zero *Wronskian*.

13 • 3. Definition

Let f and g be two differentiable functions. The *Wronskian* is the function given by

$$W[f, g] = \det\begin{bmatrix} f & g \\ f' & g' \end{bmatrix} = fg' - f'g.$$

13 • 4. Theorem

Consider the differential equation

$$y'' + p(t)y' + q(t)y = 0, \qquad y(t_0) = y_0, \qquad y'(t_0) = y'_0,$$

for some p, and q. Suppose ϕ_1 and ϕ_2 are distinct solutions to this.

So long as $W[\phi_1, \phi_2](t_0) \neq 0$, any solution y around t_0 can be written as a linear combination of ϕ_1 and ϕ_2:

$$y(t) = C\phi_1(t) + D\phi_2(t),$$

for constants C and D. ϕ_1 together with ϕ_2 make up a *fundamental set of solutions*

In some sense, we're making sure that f and g are different enough: that the vector $\langle f, f' \rangle$ isn't just a scaled version of $\langle g, g' \rangle$, and similarly that $\langle f, g \rangle$ isn't just a scaled version of $\langle f', g' \rangle$. In another sense, we're making sure that 0 isn't a linear combination of these (when the coefficients are non-zero). Note the following theorem that tells us what $W[\phi_1, \phi_2]$ is without having to find ϕ_1 and ϕ_2:

13 • 5. Theorem (Abel's Theorem)

Consider the differential equation

$$y'' + p(t)y' + q(t)y = 0,$$

and suppose $_1$ and ϕ_2 are solutions to this. Therefore

$$W[\phi_1, \phi_2](t) = Ce^{-\int p(t)\,\mathrm{d}t}$$

for some constant C (depending on ϕ_1 and ϕ_2).

In particular, this means that if the Wronskian of two solutions is 0 somewhere on an interval, since $e^{(\text{whatever})}$ is never 0, C must be 0. This would imply that it is always 0 on that interval, so that the inverse holds too: if the Wronskian is not 0 somewhere on an interval, then it is never 0 on that interval.

In the case that p and q are continuous everywhere, this means the Wronskian of two solutions is either always 0 or never 0. However, if p and q have discontinuities, the Wronskian may be non-0 just on a small interval, and 0 everywhere else. Regardless, we still get our desired equalities from previous theorems: on such an interval, any solution can be written as a linear combination of those two solutions.

Note also that Abel's Theorem (13 • 5) references the Wronskian of solutions to second-order, homogeneous, differential equations, not functions in general. For example, the Wronskian of e^t and $\cos(t)$ is

$$W[e^t, \sin t] = \det \begin{bmatrix} e^t & \sin t \\ e^t & \cos t \end{bmatrix} = e^t \cos t - e^t \sin t,$$

which is never constantly 0 on an interval[i], but it is occasionally 0, e.g. at $t = \pi/4$.

Now returning back to equations, in the case that P, Q, and R are constants in the equation $P(t)y'' + Q(t)y' + R(t)y = 0$, we have the following result that allows us to characterize the solutions.

13 • 6. Result

Consider the differential equation $ay'' + by' + cy = 0$. If there are two solutions $r_1 \neq r_2$ (real or complex) to

$$ar^2 + br + c = 0,$$

then all solutions of the differential equation are of the form

$$Ce^{r_1 t} + De^{r_2 t}$$

for some constants C and D.

A hiccup, however, is when there is only *one* solution to $ar^2 + br + c = 0$, i.e. when using the quadratic formula, the *discriminant* is 0.

13 • 7. Result

Consider the differential equation $ay'' + by' + cy = 0$. If $r = r_1 = r_2$ is the only solution to

$$ar^2 + br + c = 0,$$

then all solutions of the differential equation are of the form

$$Ce^{rt} + Dte^{rt},$$

for some constants C and D.

Proof .:.

Claim 1

$\phi_1(t) = te^{rt}$ is a solution to $ay'' + by' + cy = 0$.

[i] at least on an interval that isn't a single point

Proof .:.

We can calculate that

$$\phi_1'(t) = rte^{rt} + e^{rt}, \qquad \phi_1''(t) = r^2te^{rt} + 2re^{rt}.$$

Hence

$$\begin{aligned} a\phi_1'' + b\phi_1' + c\phi_1 &= ar^2te^{rt} + 2are^{rt} + brte^{rt} + be^{rt} + cte^{rt} \\ &= te^{rt}(ar^2 + br + c) + e^{rt}(2ar + b) \\ &= c + e^{rt}(2ar + b). \end{aligned}$$

Since r is the only solution to $ax^2 + bx + c = 0$, we can write

$$ax^2 + bx + c = a(x - r)^2 = ax^2 - 2arx + ar^2,$$

which then requires b to be $-2ar$. Hence $2ar + b = 0$ so that above,

$$a\phi_1'' + b\phi_1' + c\phi_1 = 0$$ ⊣

It's clear that $\phi_2(t) = e^{rt}$ is a solution to $ay'' + by' + cy = 0$ just by computation. So these two form two solutions. Now we want to use Theorem 13 • 4, so we must find when the Wronskian is non-zero.

$$W[\phi_1, \phi_2](t) = \det\begin{bmatrix} \phi_1 & \phi_2 \\ \phi_1' & \phi_2' \end{bmatrix} = \det\begin{bmatrix} te^{rt} & e^{rt} \\ e^{rt} + rte^{rt} & re^{rt} \end{bmatrix} = tre^{2rt} - e^{2rt} - rte^{2rt} = -e^{2rt},$$

which is never 0. So by Theorem 13 • 4, all solutions are of the form

$$y(t) = C\phi_1(t) + D\phi_2(t),$$

for constants C and D ⊣

To introduce *more* notation, when working with a homogeneous differential equation $y'' + p(t)y' + q(t)y = 0$, some people like to write $L[\phi]$ as the function on the left hand side of that equality:

$$L[\phi] = \phi'' + p\phi' + q\phi.$$

Note that the differential equation is then stating $L[y] = 0$. Obviously not all functions ϕ have $L[\phi] = 0$, so it is up to the above methods to try and figure out which ϕ satisfy this. Right now, we only know how to do this when the differential equations involves constant coefficients. But supposing that we're giving one ϕ_1, how do we find a second solution ϕ_2? Below is a method to help with this.

§ 13 A. Reduction of order

Returning to differential equations, recall that if we can find two "different enough" solutions to a second-order, homogeneous, differential equation

$$P(t)y'' + Q(t)y' + R(t)y = 0,$$

then *any* solution will be a linear combination of those two solutions: $Cy_1 + Dy_2$ for constants C and D. But often it can be hard to find two solutions, let alone one. The method of *reduction of order* can sometimes be used to help with this. In general, reduction of order is just a means of reducing the order of a differential equation by considering a different differential equation with a smaller order.

13 A • 1. Result

Consider the differential equation $y'' + p(t)y' + q(t)y = 0$. Suppose y_1 is a solution to this (that isn't just the constant 0). Thus

$$y_2(t) = v(t) \cdot y_1(t)$$

is also a solution, given that we can find a v where

$$y_1(t)v'' + (2y_1'(t) + p(t)y_1(t))v' = 0.$$

Proof .:.

If we can find such a v, then by the product rule,

$$\begin{aligned} y_2' &= v'y_1 + vy_1' \\ y_2'' &= v''y_1 + 2v'y_1' + vy_1'' \end{aligned}$$

$$
\begin{aligned}
&= v''y_1 + 2v'y_1' - vpy_1' - vqy_1. \\
\therefore\ y_2'' + py_2' + qy_2 &= (v''y_1 + 2v'y_1' - vpy_1' - vqy_1) + pv'y_1 + pvy_1' + qvy_1 \\
&= v''y_1 + (2y_1' - v + py_1)v' \\
&= 0
\end{aligned}
$$
⊣

The equation above used to find v is really just a result of plugging in $y_2 = vy_1$ to the differential equation, and finding out what v must satisfy. Actually finding a v where that happens can be done by instead solving for $w = v'$ in the first-order, linear differential equation

$$y_1w' + (2y_1 + py_1)w = 0,$$

and then integrating w to get v.

Section 14. *Non-Homogeneous Equations*

§ 14 A. Theory behind non-homogeneous equations

A homogeneous, second-order, differential equation is any equation of the form

$$P(t)y'' + Q(t)y' + R(t)y = 0.$$

We "know" how to solve a second-order, homogeneous differential equation in general—at least in principle. Doing this requires finding two particular solutions that are "different enough"—in that they have a non-zero Wronskian. Finding two solutions can be incredibly difficult, but once we have them, ϕ_1 and ϕ_2 to

$$P(t)y'' + Q(t)y' + R(t)y = 0,$$

then *all* solutions to this will be of the form

$$y(t) = C\phi_1(t) + D\phi_2(t)$$

for various constants C and D.

Now we will turn to non-homogeneous, linear, second-order, differential equations.

14 A • 1. Definition

A non-homogeneous, linear, second-order, differential equation is an equation of the form

$$P(t)y'' + Q(t)y' + R(t)y = S(t)$$

for functions P, Q, R, and S.

Solving these is even harder in general, so we will look at a couple of methods to help. The guiding idea is the following result.

14 A • 2. Result

Consider the homogeneous differential equation

$$P(t)y'' + Q(t)y' + R(t)y = 0.$$

Let ϕ_1 and ϕ_2 be solutions of this with non-zero Wronskian. Thus all solutions to the *non-homogeneous* differential equation

$$P(t)y'' + Q(t)y' + R(t)y = S(t)$$

will be of the form

$$y(t) = C\phi_1(t) + D\phi_2(t) + Y(t),$$

where $Y(t)$ solves the non-homogeneous equation.

Proof ∴.

Firstly, let $y(t)$ be an arbitrary solution to the non-homogeneous, differential equation. Consider the function $\phi = y - Y$. We have derivatives $\phi' = y' - Y'$, and $\phi'' = y'' - Y''$. Hence

$$\begin{aligned} P\phi'' + Q\phi' + R\phi &= P \cdot (y'' - Y'') + Q \cdot (y' - Y') + R \cdot (y - Y) \\ &= (Py'' + Qy' + Ry) - (Py'' + QY' + RY) \\ &= S - S = 0. \end{aligned}$$

In other words, ϕ is a solution to the homogeneous differential equation

$$P(t)y'' + Q(t)y' + R(t)y = 0.$$

But those solutions are of the form $C\phi_1 + D\phi_2$ for some constants C and D. Thus for some constants C, D,

$$\begin{aligned} y - Y &= C\phi_1 + D\phi_2 \\ \therefore y &= C\phi_1 + D\phi_2 + Y \end{aligned}$$

⊣

So if solving a second-order, homogeneous, differential equation requires finding two solutions—ϕ_1 and ϕ_2—solving a non-homogeneous, linear, differential equation requires finding three—ϕ_1 and ϕ_2 to the homogeneous version, and Y to the non-homogeneous version. The next two ideas are trying to find such a Y.

§ 14 B. Undetermined coefficients

The method of undetermined coefficients is in essence making an educated guess at a particular solution Y to an equation of the form

$$ay'' + by' + cy = g(t)$$

for various $g(t)$. Note that this is somewhat restricted for two reasons. Firstly, rather than second-order, linear, differential equations in general,

$$P(t)y'' + Q(t)y' + R(t) = g(t),$$

we require that the coefficients P, Q, and R be constants. Secondly, g must be simple or nice enough to be able to guess the solution. This means products and sums of polynomials, sines, cosines, and exponentials. If g involves a quotient, or some other function, we can't use the method of undetermined coefficients.

14 B • 1. Result (Undetermined Coefficients)

Consider the differential equation $ay'' + by' + cy = g(t)$.

If $g(t) = a_n t^n + \cdots + a_1 t^1 + a_0$,	then guess	$Y(t) = A_n t^n + \cdots + A_1 t^1 + A_0$;
if $g(t) = \sin(at)$, or $g(t) = \cos(at)$,	then guess	$Y(t) = A\sin(at) + B\cos(at)$;
if $g(t) = e^{at}$,	then guess	$Y(t) = Ae^{at}$;

for various constants.

If the guess ends up being a solution to the homogeneous equation $ay'' + by' + cy = 0$, then just multiply the guess by t. And if that also is a solution, keep multiplying by t until it isn't.

If g is a product/sum of these, then our guess should be a product/sum of the corresponding guesses. For example, if $g(t) = t^3 \sin(at)e^{bt}$, we should guess

$$Y(t) = (C_3t^3 + C_2t^2 + C_1t + C_0) \cdot (A\sin(at) + B\cos(at)) \cdot Ge^{bt},$$

for undetermined coefficients C_0, C_1, C_2, C_3, A, B, and G. Note we need different constants for each guess involved.

Noting the restrictions on the method of undetermined coefficients is where variation of parameters comes in.

§ 14 C. Variation of parameters

A better method, but harder to actually do, is variation of parameters. Although significantly more general, applying to differential equations of the form

$$y'' + p(t)y' + q(t)y = g(t),$$

the method is harder to work with, and harder to remember. The basic thing to remember is the setup $Y(t) = u_1\phi_1 + u_2\phi_2$ for solutions ϕ_1, and ϕ_2 to the homogeneous version, and

$$\begin{bmatrix} \phi_1 & \phi_2 \\ \phi_1' & \phi_2' \end{bmatrix} \cdot \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} 0 \\ g \end{bmatrix}.$$

You can then solve for u_1 and u_2 to find Y. This will generalize for higher-order differential equations, when we have the setup that $Y = u_1\phi_1 + \cdots + u_n\phi_n$, and then assume

$$\mathsf{W} \cdot \begin{bmatrix} u_1' \\ \vdots \\ u_{n-1}' \\ u_n' \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ g \end{bmatrix},$$

where W is the Wronskian matrix of $\phi_1, \cdots, \phi_n$. So it's good to memorize the first matrix equation.

Using either method, once we have ϕ_1, ϕ_2—fundamental solutions to the homogeneous version of the differential equation—and a Y, the general solution to the non-homogeneous equation will be

$$C\phi_1 + D\phi_2 + Y$$

for constants C and D.

14 C • 1. Result (Variation of Parameters)

Consider the differential equation

$$y'' + p(t)y' + q(t)y = g(t).$$

Let ϕ_1 and ϕ_2 be two solutions to the homogeneous version $y'' + p(t)y' + q(t)y = 0$ with $W[\phi_1, \phi_2] \neq 0$. Thus the following Y is a solution to the non-homogeneous equation.

$$Y = -\phi_1 \int \frac{\phi_2 g}{W[\phi_1, \phi_2]}\,\mathrm{d}t + \phi_2 \int \frac{\phi_1 g}{W[\phi_1, \phi_2]}\,\mathrm{d}t.$$

So all solutions are of the form $y = C\phi_1 + D\phi_2 + Y$ for constants C and D.

Proof .:.

We begin by assuming two things:

$$Y(t) = u_1(t)\phi_1(t) + u_2(t)\phi_2(t)$$
$$Y'(t) = u_1(t)\phi_1'(t) + u_2(t)\phi_2'(t),$$

for some functions u_1 and u_2. This ensures that u_1 and u_2 act *similar* to constants when we take the first derivative. The second derivative, however, allows for that to change:

$$\begin{aligned} Y'' &= \frac{\mathrm{d}}{\mathrm{d}t}(u_1\phi_1') + \frac{\mathrm{d}}{\mathrm{d}t}(u_2\phi_2') \\ &= u_1'\phi_1' + u_1\phi_1'' + u_2'\phi_2' + u_2\phi_2''. \end{aligned}$$

Note that then Y is a solution to the given differential equation iff

$$\begin{aligned} Y'' + pY' + qY = g(t) &\longleftrightarrow (u_1'\phi_1' + u_1\phi_1'' + u_2'\phi_2' + u_2\phi_2'') + p(u_1\phi_1' + u_2\phi_2') + q(u_1\phi_1 + u_2\phi_2) = g(t) \\ &\longleftrightarrow u_1'\phi_1' + u_1(\phi_1'' + p\phi_1' + q\phi_1) + u_2'\phi_2' + u_2(\phi_2'' + p\phi_2' + q\phi_2) = g(t) \\ &\longleftrightarrow u_1'\phi_1' + u_2'\phi_2' = g(t). \end{aligned}$$

Now because $Y'(t) = u_1\phi_1' + u_2\phi_2'$, the rest of the terms we would get from the product rule must be 0:

$$u_1'\phi_1 + u_2'\phi_2 = 0.$$

Hence we have two equations, and two unknowns—u_1' and u_2'—which is equivalent to the matrix equation

$$\begin{bmatrix} \phi_1 & \phi_2 \\ \phi_1' & \phi_2' \end{bmatrix} \cdot \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} 0 \\ g \end{bmatrix}.$$

Since the Wronskian—the determinant of the matrix on the left—is non-zero, we can invert the matrix to get the equivalent statement that

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \frac{1}{W[\phi_1, \phi_2]} \begin{bmatrix} \phi_2' & -\phi_2 \\ -\phi_1' & \phi_1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ g \end{bmatrix} = \frac{1}{W[\phi_1, \phi_2]} \begin{bmatrix} -\phi_2 g \\ \phi_1 g \end{bmatrix}.$$

So any functions u_1 and u_2 satisfying this will have $Y = u_1\phi_1 + u_2\phi_2$ satisfy the differential equation. In particular, integrating tells us that

$$u_1 = \int \frac{-\phi_2 g}{W[\phi_1, \phi_2]}\, \mathrm{d}t, \qquad \text{and} \qquad u_2 = \int \frac{\phi_1 g}{W[\phi_1, \phi_2]}\, \mathrm{d}t$$

work. Thus we have found a particular solution Y, and so by Result 14 A • 2, any solution will be of the form $C\phi_1 + D\phi_2 + Y$ for constants C and D. ⊣

So in general, the process is to assume

$$Y = u_1\phi_1 + u_2\phi_2$$

where

$$u_1'\phi_1 + u_2'\phi_2 = 0$$
$$u_1'\phi_1' + u_2'\phi_2' = g,$$

and then to solve for u_1' and u_2' to get some u_1 and u_2 that work.

Section 15. *Exercises*

IV • Ex1. Exercise: Find the general solution to $y'' + 2y' + 2y = 0$.

IV • Ex2. Exercise: Consider the initial value problem

$$3u'' - u' + 2u = 0, \qquad u(0) = 2, \quad u'(0) = 0.$$

1. Find the solution $u(t)$ of this problem.
2. For $t > 0$, find the first time at which $|u(t)| = 10$.

IV • Ex3. Exercise: Consider the initial value problem

$$5u'' + 2u' + 7u = 0, \qquad u(0) = 2, \quad u'(0) = 1.$$

1. Find the solution $u(t)$ of this problem.
2. Find the smallest T such that $|u(t)| \le 0.1$ for all $t > T$.

IV • Ex4. Exercise: Show that $W[e^{\lambda t}\cos(\mu t), e^{\lambda t}\sin(\mu t)] = \mu e^{2\lambda t}$.

IV • Ex5. Exercise: An equation of the form

$$t^2 \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \alpha t \frac{\mathrm{d}y}{\mathrm{d}t} + \beta y = 0, \quad t > 0, \tag{33}$$

where α and β are real constants, is called an *Euler equation*.

a. Let $x = \ln t$ and calculate $\frac{\mathrm{d}y}{\mathrm{d}t}$ and $\frac{\mathrm{d}^2 y}{\mathrm{d}t^2}$ in terms of $\frac{\mathrm{d}y}{\mathrm{d}x}$ and $\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}$.
b. Use the results of part (a) to transform equation (33) into

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + (\alpha - 1)\frac{\mathrm{d}y}{\mathrm{d}x} + \beta y = 0. \tag{34}$$

Observe that differential equation (34) has constant coefficients. If $y_1(x)$ and $y_2(x)$ form a fundamental set of solutions of equation (34), then $y_1(\ln t)$ and $y_2(\ln t)$ form a fundamental set of solutions of equation (33).

IV • Ex6. Exercise: Find the general solution of $y'' - 2y' + y = 0$.

IV • Ex7. Exercise: Find the general solution of $16y'' + 24y' + 9y = 0$.

IV • Ex8. Exercise: Solve the initial value problem

$$y'' + 4y' + 4y = 0, \qquad y(-1) = 2, \quad y'(-1) = 1.$$

Sketch the graph of the solution and describe its behavior for increasing t.

IV • Ex9. Exercise: Solve the equation

$$y'' + y' + 3y = 0$$

with initial conditions $y(0) = 1$, and $y'(0) = 1$.

IV • Ex10. Exercise:

a Consider the equation $y'' + 2ay' + a^2 y = 0$. Show that the roots of the characteristic equation are $r_1 = r_2 = -a$ so that one solution of the equation is e^{-at}.

b Use Abel's theorem [$W(t) = ce^{\int -p(t)\,\mathrm{d}t}$ for c a constant, and $p(t)$ in $y'' + p(t)y' + q(t)y = 0$] to show that the Wronskian of any two solutions of the given equation is

$$W(t) = y_1(t)y_2'(t) - y_1'(t)y_2(t) = c_1 e^{-2at},$$

where c_1 is a constant.

c Let $y_2(t) = e^{-at}$ and use the result of part (b) to obtain a differential equation satisfied by a second solution $y_2(t)$. By solving this equation, show that $y_2(t) = te^{-at}$.

IV • Ex11. Exercise: Use the method of reduction of order to find a second solution of the differential equation

$$t^2 y'' + 2ty' - 2y = 0, \quad t > 0; \qquad y_1(t) = t.$$

IV • Ex12. Exercise: The differential equation

$$y'' + \delta(xy' + y) = 0$$

arises in the study of turbulent flow of a uniform stream past a circular cylinder.

1. Verify that $y_1(x) = \exp(-\delta x^2/2)$ is one solution, and then
2. find the general solution in the form of an integral.

IV • Ex13. Exercise: Consider the differential equation

$$y'' + y = 3 + 4\sin(2t).$$

Give the general solution.

IV • Ex14. Exercise: Consider the differential equation

$$y'' + y = \tan t.$$

Find the general solution.

IV • Ex15. Exercise: Find the general solution to $y'' - 2y' - 3y = -3te^{-t}$.

IV • Ex16. Exercise: Find the general solution to $y'' + y = 3\sin(2t) + t\cos(2t)$.

IV • Ex17. Exercise: Find the general solution to

$$y'' + y' + 4y = 2\sinh t$$

given that $\sinh t = (e^t - e^{-t})/2$.

IV • Ex18. Exercise: Solve the equation

$$y'' + 2y' + 5y = t + 2\sin t,$$

with initial conditions $y(0) = 1$, and $y'(0) = 0$.

IV • Ex19. Exercise: Consider the equation $y'' - 3y' - 4y = 2e^{-t}$ with

$$\phi_1(t) = e^{-t}, \qquad \text{and} \qquad \phi_2(t) = e^{4t}$$

solutions to the homogeneous version. Seek a solution to the non-homogeneous equation of the form

$$Y(t) = v(t) \cdot \phi_1(t) = v(t)e^{-t}$$

for some v.

a. Substitute Y, Y', and Y'' into the non-homogeneous equation and show that v must satisfy $v'' - 5v' = 2$.
b. Let $w = v'$ and show that w satisfies $w' - 5w = 2$, solving this for w.
c. Integrate w to find v and then show that

$$Y = -\frac{2}{5}te^{-t} + \frac{1}{5}c_1e^{4t} + c_2e^{-t}.$$

The first term on the right-hand side is the desired particular solution of the nonhomogeneous equation. Note that it is a product of t and e^{-t}.

IV • Ex20. Exercise: Consider $ay'' + by' + cy = g(t)$ where a, b, c are positive.

Let Y_1 and Y_2 be solutions to this. Show that $Y_1 - Y_2 \to 0$ as $t \to \infty$. Is this true if $b = 0$?

IV • Ex21. Exercise: Consider the equation $y'' + by' + cy = g(t)$, and suppose r_1 and r_2 (not necessarily distinct nor real) are zeroes of the characteristic polynomial of the corresponding homogeneous equation.

1. Verify that the non-homogeneous differential equation can be written as

$$(D - r_1)(D - r_2)y = g(t)$$

where $r_1 + r_2 = -b$ and $r_1r_2 = c$.
2. Let $u = (D - r_2)y$. Solve $y'' + by' + cy = g(t)$ by solving the two equations

$$(D - r_1)u = g(t), \qquad (D - r_2)y = u(t).$$

IV • Ex22. Exercise: Use the method of variation of parameters to find a particular solution to

$$y'' - 5y' + 6y = 2e^t$$

IV • Ex23. Exercise: Find the general solution to

$$4y'' + y = 2\sec(t/2), \qquad -\pi < t < \pi.$$

IV • Ex24. Exercise: Consider the differential equation, for $x > 0$,

$$x^2y'' + xy' + \left(x^2 - \frac{1}{4}\right)y = g(x)$$

for g an arbitrary continuous function.

1. Show

$$\phi_1(x) = \frac{1}{\sqrt{x}}\sin x, \qquad \text{and} \qquad \phi_2(x) = \frac{1}{\sqrt{x}}\cos x$$

are solutions to the homogeneous version.
2. Find the general solution to the non-homogeneous equation.

IV • Ex25. Exercise: Solve the equation $y'' + y' + 10y = 0$; with $y(0) = 1$, $y'(0) = 0$.

IV • Ex26. Exercise: Solve the initial value problem

$$y'' + y' + 5y = t - e^{-t},$$

where $y(0) = 0$, and $y'(0) = 0$.

IV • Ex27. Exercise: Solve the equation

$$y'' + 3y = \begin{cases} t & \text{if } 0 \le t \le 1 \\ 1 & \text{if } 1 < t, \end{cases}$$

where $y(0) = 1$, and $y'(0) = 1$.

There is another way to do this problem, which is by undetermined coefficients. To do this, however, we would need to do it in an unusual way. The homogeneous solution, rather than $C\phi_1 + D\phi_2$ like usual, needs to have separate constants over each interval. In essence, we must solve the equation

$$y'' + 3y = t$$

and then the equation

$$y'' + 3y = 1,$$

and then glue the two solutions together at $t = 1$. The next solution to the same problem gives this method.

IV • Ex28. Exercise: Solve the equation

$$y'' + 3y = \begin{cases} t & \text{if } 0 \le t < 1 \\ 1 & \text{if } 1 < t, \end{cases}$$

where $y(0) = 1$, and $y'(0) = 1$.

IV • Ex29. Exercise: Solve the equation

$$x^2 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 2x \frac{\mathrm{d}y}{\mathrm{d}x} + 3y = 0$$

with initial conditions $y(2) = 1$ and $y'(2) = 0$ for $x \ge 2$.

Chapter V. Using Linear Algebra

Section 16. *Gaussian Elimination*

§ 16 A. Motivating Gaussian elimination

Our goal here is just to go over *Gaussian elimination*, a process to invert a matrix. Our setup will be equations of the form $\mathsf{A}\vec{v} = \vec{w}$ for a given matrix A, and vectors $\vec{v}$ and $\vec{w}$. In principle, we would like to say $\vec{v} = \mathsf{A}^{-1}\vec{w}$, but not all matrices can be inverted, and even when they can, finding the inverse is sometimes a difficult task. Gaussian elimination is a procedure to find $\vec{v}$ given $\vec{w}$ and A, and in essence, find such an inverse if it exists.

This process can be useful for us in differential equations when looking at higher-order differential equations—i.e. third-order, fourth-order, and so on. In particular, linear, non-homogeneous, differential equations require us to solve the matrix equation

$$\mathsf{W} \cdot \begin{bmatrix} u_1' \\ \vdots \\ u_{n-1}' \\ u_n' \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_n \\ \phi_1' & \phi_2' & \cdots & \phi_n' \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1^{(n-1)} & \phi_2^{(n-1)} & \cdots & \phi_n^{(n-1)} \end{bmatrix} \cdot \begin{bmatrix} u_1' \\ \vdots \\ u_{n-1}' \\ u_n' \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ g \end{bmatrix},$$

where W is the Wronskian matrix. Solving this gives $Y = u_1\phi_1 + \cdots + u_n\phi_n$ as a particular solution: the n^{th}-order version of variation of parameters. Things get more complicated especially when we consider systems of differential equations. For more information about this, consider Chapter VI.

§ 16 B. Motivation from systems of equations

Recall that with just a couple linear equations

$$\begin{aligned} Ax + By &= N \\ Cx + Dy &= M, \end{aligned}$$

we can solve the equation by adding and subtracting the respective equalities from each other: consider the following as an example.

16 B • 1. Example

Consider the system of equations

$$\begin{aligned} -x + 3y &= 5 \\ 2x + 5y &= 10. \end{aligned}$$

Procedure .:.

Adding "twice the first equation" to the second equation means that we can conclude

$$\begin{aligned} (2x + 5y) + 2(-x + 3y) &= 10 + 2 \cdot 5 \\ 11y = 20 &\longrightarrow y = 20/11. \end{aligned}$$

Using this result, the first equation then says $-x + 60/11 = 5$, i.e. $x = 5/11$.

This can allow us to quickly solve the equations if the coefficients are nice enough like here. Gaussian elimination is just reformulating this idea in the context of matrices.

Matrix Procedure .:.

We could instead write the setup for this example as the matrix equation

$$\begin{bmatrix} -1 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}.$$

the idea of "adding twice the first equation to the second" just means multiplying by the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

This transforms our matrix equation into

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$
$$\begin{bmatrix} -1 & 3 \\ 0 & 11 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 20 \end{bmatrix},$$

which expresses the same equality as before: $11y = 20$. We can then muliply by a scaling matrix, and so yield the equality below.

$$\begin{bmatrix} -1 & 0 \\ 0 & 1/11 \end{bmatrix} \cdot \begin{bmatrix} -1 & 3 \\ 0 & 11 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1/11 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 20 \end{bmatrix}$$
$$\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -5 \\ 20/11 \end{bmatrix},$$

adding thrice the second row to the first then yields

$$\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -5 \\ 20/11 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5/11 \\ 20/11 \end{bmatrix},$$

In other words, $x = 5/11$, and $y = 20/11$.

This has just been motivation, however. We still need to explicitly lay out the process.

§ 16 C. Setting up Gaussian elimination

Our goal is to manipulate the matrix equation

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \quad \text{into} \quad \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

for some $c_1, \cdots, c_n$. Note that this equation on the right is just saying $\mathrm{x} = \vec{c}$. We do this by applying three operations multiple times. All of these operations are just multiplying both sides of the equation by an invertible matrix on the left.

16 C • 1. Definition

An *elementary row operation* on a matrix is an operation that

1. scales a row;
2. adds (a multiple of) one row to another; or
3. switches two rows.

Note that these operations can be accomplished by multiplying by a certain invertible matrix on the left. This allows us to transform matrices into other matrices which are much nicer to work with. In particular, we will care about *reduced row echelon form*.

16 C • 2. Definition

An entry a in a matrix A is a *leading coefficient* iff it is the first[i] non-0 entry in its row of A.

A matrix A is in *reduced row echelon form* iff

1. every leading coefficient of A is 1;
2. every leading coefficient has only 0s in its column;
3. every row's leading coefficient appears to the right of higher rows'; and
4. rows with only 0s appear at the bottom.

All the following are examples of matrices in reduced row echelon form ('$\star$' just represents some entry that doesn't need to be 0)

$$\checkmark \quad \begin{bmatrix} 1 & \star & 0 & \star & \star & 0 \\ 0 & 0 & 1 & \star & \star & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & \star & \star \\ 0 & 1 & \star & \star \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

All the following are examples of matrices that *are not* in reduced row echelon form.

$$\times \quad \begin{bmatrix} 1 & \star & \star & \star & \star & 0 \\ 0 & 1 & \star & \star & \star & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & \star & \star \\ 1 & 0 & \star & \star \end{bmatrix}, \quad \begin{bmatrix} 1 & \star & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The first isn't since row 2's leading coefficient has a non-0 entry right above it. The second isn't because the second row's leading coefficient appears to the left of the first row's. The third isn't because the there is a row of 0s which doesn't appear at the bottom.

This is all the background we need to get actually define the method of Gaussian elimination.

§ 16 D. The method of Gaussian elimination

So this is the setup to Gaussian elimination: we have a matrix equation $Ax = b$, and we want to transform A into reduced row echelon form by way of elementary row operations.

16 D • 1. Definition

Consider the equation $A \cdot x = b$ for a given matrix A, given vector b, and vector of variables x.

Gaussian elimination is the process of transforming the matrix $[A|b]$—just the entries of A next to the entries of b—by applying elementary row operations into reduced row echelon form.

Often it's useful to instead do the process on a different matrix that combines A and b. Looking back to Example 16 B • 1, we can again go through the process, starting with

$$\left[\begin{array}{cc|c} -1 & 3 & 5 \\ 2 & 5 & 10 \end{array}\right]$$

and doing the same operations as before:

$$\left[\begin{array}{cc|c} -1 & 3 & 5 \\ 2 & 5 & 10 \end{array}\right] \mapsto \left[\begin{array}{cc|c} -1 & 3 & 5 \\ 0 & 11 & 20 \end{array}\right] \mapsto \left[\begin{array}{cc|c} 1 & -3 & -5 \\ 0 & 1 & 20/11 \end{array}\right] \mapsto \left[\begin{array}{cc|c} 1 & 0 & 5/11 \\ 0 & 1 & 20/11 \end{array}\right].$$

Very useful is the following result, which tells us that—if we can invert the matrix—Gaussian elimination takes us all the way to the answer we want: it solves the equation.

16 D • 2. Result (Gaussian elimination on invertible matrices)

Let A be an invertible matrix, and b a vector. Consider the equation $A \cdot x = b$, where x is a vector of variables.
Let R be the result of doing Gaussian elimination to A, and $\vec{c}$ the result of this process applied to b.
Thus R is the identity matrix, and $x = \vec{c}$.

[i] going from left to right

Proof .:.

It turns out (unproven here) that the reduced row echelon form of a matrix is unique in the sense that there is only one matrix P such that some non-zero matrix M has $M \cdot \mathsf{A} = P$ in reduced row echelon form.

Given that A is invertible, $\mathsf{A}^{-1} \cdot \mathsf{A} = \mathrm{Id}$ is in reduced row echelon form, and so $\mathsf{R} = \mathrm{Id}$. So the process of doing Gaussian elimination is the same as applying A^{-1}. Thus $\vec{c} = \mathsf{A}^{-1}\mathsf{b} = \mathsf{x}$. ⊣

If we can invert the matrix, we have a pretty simple idea to solve the equation. The idea generalizes easily from the 3×3 case. Suppose we have the augmented matrix

$$\left[\begin{array}{ccc|c} \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{array}\right].$$

First, we just scale the first row so that the first entry becomes 1:

$$\left[\begin{array}{ccc|c} 1 & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{array}\right].$$

Now we subtract the appropriate amount of row 1 from the other two rows to clear out the first column:

$$\left[\begin{array}{ccc|c} 1 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \end{array}\right].$$

Now we do the same thing for the second column, working with the second row:

$$\left[\begin{array}{ccc|c} 1 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \end{array}\right] \mapsto \left[\begin{array}{ccc|c} 1 & \star & \star & \star \\ 0 & 1 & \star & \star \\ 0 & \star & \star & \star \end{array}\right] \mapsto \left[\begin{array}{ccc|c} 1 & 0 & \star & \star \\ 0 & 1 & \star & \star \\ 0 & 0 & \star & \star \end{array}\right],$$

and then the same for the third. The general process is to scale to 1, and then clear out the rest of the column. This allows us to work with the other columns more easily. In a similar way, switching rows can be helpful, like if there is already a 1 in the first column. Switching that row and the first row would put a 1 in the top left, and so you wouldn't need to scale. Similarly, if just through a simple addition/subtraction of a row, we can get a 1, that may be easier than scaling. Ultimately, we'll arrive at the same answer, but the computations involved are already tedious, and involving fractions can make it more-so, so there's a bit of a disincentive to scale if we don't have to. We also don't need to start with the first column: we can do them in any order.

Note that we can also generalize this process to find the inverse of a matrix directly (if there is an inverse). In particular, we can find the first column of the inverse matrix by solving

$$\mathsf{A} \cdot \mathsf{x} = [1\ 0\ \cdots\ 0]^\top.$$

Similarly, we can find the second column of the inverse matrix, and then the third, fourth, etc. through the same sort of process, solving for x in each of these. But note that we can absorb this iterative process into just one process, solving

$$\mathsf{A} \cdot \mathsf{X} = \mathrm{Id}\,.$$

So instead of starting with many augmented matrices $[\mathsf{A} \mid e]$ for some vector e—a bunch of 0s and one 1—we can start with the augmented matrix below, and go as before to turn this into reduced row echelon form; turning A into Id.

$$[\mathsf{A} \mid \mathrm{Id}] = \left[\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & 1 & & 0 \\ \vdots & \ddots & \vdots & & \ddots & \\ a_{n1} & \cdots & a_{nn} & 0 & & 1 \end{array}\right].$$

Section 17. *Matrices*

There are a few fundamental operations on matrices we will be looking at, namely multiplication, addition, scaling, transposition, and conjugation. These are defined as follows.

17 • 1. Definition

onsider the matrix A, where the i, jth slot is a_{ij}. This is written $A = (a_{ij})_{i,j}$. Here i refers to the row, and j to the column.

Consider the matrices $A = (a_{ij})_{i,j}$, and $B = (b_{ij})_{i,j}$. Let k be a constant.

1. $A + B = (a_{ij} + b_{ij})_{i,j}$, i.e. the matrix where we add the corresponding entries.
2. $kA = (ka_{ij})_{i,j}$, i.e. the matrix where we scale all the entries.
3. $A \cdot B$ is given by the following:
$$A \cdot B = \left(\sum_k a_{ik} b_{kj}\right)_{i,j}.$$
4. $A^\top$, *the transpose*, is $(a_{ji})_{i,j}$.
5. $\overline{A}$, *the conjugate*, is $(\overline{a_{ij}})_{i,j}$, taking the complex conjugate of all the entries.
6. $A^* = \overline{(A^\top)} = (\overline{A})^\top$, taking the conjugate and transposing (in either order).

Matrix multiplication can be understood as successive multiplication of columns and rows. Consider the equality
$$\begin{bmatrix} a & b & c \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \end{bmatrix}.$$
We can then define matrix multiplication of a full matrix by doing this row by row on the left:
$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + jz \end{bmatrix}.$$
Now we can define full matrix multiplication by doing this column by column on the right
$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix} \cdot \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} = \begin{bmatrix} ax_1 + by_1 + cz_1 & ax_2 + by_2 + cz_2 & ax_3 + by_3 + cz_3 \\ dx_1 + ey_1 + fz_1 & dx_2 + ey_2 + fz_2 & dx_3 + ey_3 + fz_3 \\ gx_1 + hy_1 + jz_1 & gx_2 + hy_2 + jz_2 & gx_3 + hy_3 + jz_3 \end{bmatrix}.$$
Matrix multiplication can be understood as composition, though this isn't so important for us. For example, suppose $a = 2x - y$ and $b = 5x + y$, and $v = a + b$, $w = -2a - b$. These are represented by the matrix equations
$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 5 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix}.$$
If we wanted to write v and w in terms of x and y instead of a and b, we would get
$$v = (2x - y) + (5x + y) = 7x + 0y, \qquad w = -2(2x - y) - (5x + y) = -9x + y,$$
which, when represented as a matrix equation, is precisely what we get when we multiply the matrices:
$$\begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ 5 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 & 0 \\ -9 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}.$$
Inverting a matrix is also a kind of operation, but only exists for certain matrices, namely those which are *invertible*. This has been covered in Section 16, where Gaussian elimination is a method to find the inverse of a matrix.

We have one last definition for looking at matrices: the determinant. The determinant is very useful mostly for its properties, which actually determine the function.[i] Note also that the determinant is only defined for square matrices. Below are just some properties of the determinant.

17 • 2. Result

For square matrices A and B, of size $n \times n$,
$$\det(\mathrm{Id}) = 1$$
$$\det(A) = a_{11} \cdot a_{22} \cdot a_{33} \cdot \ldots \cdot a_{nn}, \text{ if A is triangular,}$$

[i]just like how integration on continuous functions is uniquely defined by agreeing with the area of rectangles, and adding over disjoint intervals

$$\det(\mathsf{A}^\top) = \det(\mathsf{A})$$
$$\det(\mathsf{A}\cdot\mathsf{B}) = \det(\mathsf{A})\cdot\det(\mathsf{B})$$
$$\det(c\mathsf{A}) = c^n\mathsf{A}$$
$$\det(\mathsf{A}) = 0 \text{ if two rows or columns of A are the same.}$$

The precise definition of the determinant is given below.

17 • 3. Definition

Let $\mathsf{A} = (a_{ij})_{i,j}$ be an $n \times n$ matrix. The *determinant* of A is given by

$$\det(\mathsf{A}) = \sum_{i_k, j_k = 1}^{n} \pm a_{i_1,j_1} \cdots a_{i_n,j_n},$$

where the sign of $\pm$ is given by the indices. In other words, we sum up all products of elements, where only one element from each row and column is chosen at a time, and either add or subtract this.

This hasn't actually defined the determinant completely, so the following are the determinants of 2×2 and 3×3 matrices in general.

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$
$$\det\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = a_1(b_2c_3 - c_2b_3) + b_1(c_2a_3 - a_2c_3) + c_1(a_2b_3 - b_2a_3).$$

Section 18. *Linear Algebra*

§ 18 A. Linear independence

First we have the notion of *linear independence*.[ii]

18 A • 1. Definition

Let $\mathsf{v}_0, \mathsf{v}_1, \cdots, \mathsf{v}_n$ be vectors, and $c_1, \cdots, c_n$ be constants which aren't all just 0. Consider the equation

$$c_1\mathsf{v}_0 + \cdots + c_n\mathsf{v}_n = \vec{0}.$$

If there are $c_1, \cdots, c_n$where this is true—except trivially when all are 0—then the vectors $\mathsf{v}_0, \mathsf{v}_1, \cdots, \mathsf{v}_n$ are said to be *linearly dependent*.

If this is always false—there are no $c_1, \cdots, c_n$ except trivially when all are 0—then $\mathsf{v}_0, \mathsf{v}_1, \cdots, \mathsf{v}_n$ are said to be *linearly independent*.

Being linearly independent just means you can't write one of the vectors in terms of the others by adding and scaling. Note that this is a property of *collections* of vectors, not the vectors themselves. $\langle 1, 0\rangle$ and $\langle 0, 1\rangle$ are linearly independent, but $\langle 1, 0\rangle, \langle 0, 1\rangle, \langle 1, 1\rangle$ are linearly dependent, since

$$\langle 1, 1\rangle - \langle 1, 0\rangle - \langle 0, 1\rangle = \langle 0, 0\rangle.$$

Testing linear independence can be done in a variety of ways. Note that the fundamental equation that we care about is

$$c_1\mathsf{v}_0 + \cdots + c_n\mathsf{v}_n = \vec{0}.$$

[ii] Almost everything in this section is discussing finite dimensional vector spaces. Generalizing to infinite dimensional ones requires slight changes to all this terminology in the way that I've written it here.

If we regard the vectors v_i as *column vectors*, this is just the matrix equation

$$\begin{bmatrix} | & & | \\ \mathsf{v}_0 & \cdots & \mathsf{v}_n \\ | & & | \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

On the left, each column is given by the v_i vectors, you might call each column the transpose of the vector.

18 A • 2. Result

Let A be a matrix. Thus $\det(\mathsf{A}) = 0$ iff $\mathsf{A}\mathsf{x} = \vec{0}$ has a non-$\vec{0}$ solution.

So to determine whether vectors are linearly independent, it's enough to either calculate the determinant of the matrix of column vectors $[\mathsf{v}_0 \ \cdots \ \mathsf{v}_n]$:

$$\det[\mathsf{v}_0 \ \cdots \ \mathsf{v}_n] = 0 \longleftrightarrow \text{ there is some } c_0, \cdots, c_n \text{ so that the above equality holds.}$$

Alternatively, we can proceed by Gaussian elimination to attempt to find such $c_0, \cdots, c_n$. If we end up only with $c_0 = \cdots = c_n = 0$, then we know that the vectors are linearly independent. Consider the following tests we then have.

18 A • 3. Result

A set of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is *linearly independent* iff one of the three equivalent conditions holds:

1. $\det \begin{bmatrix} — & \vec{v}_1 & — \\ & \vdots & \\ — & \vec{v}_n & — \end{bmatrix} \neq 0;$
2. $\det \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} \neq 0;$
3. $\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} \vec{c} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ only has the solution $\vec{c} = \vec{0}$.

We have the same sort of characterization for linear dependence.

18 A • 4. Result

A set of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is *linearly dependent* iff one of the three equivalent conditions holds:

1. $\det \begin{bmatrix} — & \vec{v}_1 & — \\ & \vdots & \\ — & \vec{v}_n & — \end{bmatrix} = 0;$
2. $\det \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} = 0;$
3. $\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} \vec{c} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ has a non-$\vec{0}$ solution for $\vec{c}$.

If we want an actual relationship between $\vec{v}_1, \ldots, \vec{v}_n$, the solution of the third condition gives us one:

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}.$$

§ 18 B. Eigenvalues and eigenvectors

An *eigenvector* is a vector which is only scaled by a given matrix. An *eigenvalue* is how much that vector is scaled by.

18 B • 1. Definition

Let A be a matrix. Consider the equation for $\mathsf{v} \neq \vec{0}$

$$\mathsf{A} \cdot \mathsf{v} = \lambda \mathsf{v}.$$

λ is an *eigenvalue* for A. v is an eigenvector for A corresponding to the eigenvalue λ.

Eigenvalues and eigenvectors will prove to be important for us later. Finding them can be a tricky process.

18 B • 2. Result

Let A be a square matrix. Thus λ is an eigenvalue iff

$$\det(\mathsf{A} - \lambda \, \mathrm{Id}) = 0.$$

v is an eigenvector iff there is some eigenvalue λ_i where

$$(\mathsf{A} - \lambda_i \, \mathrm{Id}) \cdot \mathsf{v} = \vec{0}.$$

Proof .:.

λ is an eigenvalue iff there is some non-$\vec{0}$ v where $\mathsf{A} \cdot \mathsf{v} = \lambda \mathsf{v}$. This is equivalent to

$$\vec{0} = \mathsf{A}\mathsf{v} - \lambda \mathsf{v} = (\mathsf{A} - \lambda \, \mathrm{Id})\mathsf{v}.$$

By Result 18 A • 2, this happens iff $\det(\mathsf{A} - \lambda \, \mathrm{Id}) = 0$.

v is an eigenvector iff there is an eigenvalue where $\mathsf{A} = \lambda_i \mathsf{v}$. Moving to the other side and factoring yields that this is equivalent to $(\mathsf{A} - \lambda_i \, \mathrm{Id})\mathsf{v} = \vec{0}$. ⊣

So we just need to solve the polynomial equation $\det(\mathsf{A} - \lambda \, \mathrm{Id}) = 0$ for various λ. Now if we know all the eigenvalues, λ_1, λ_2, ..., and λ_n, we can find the eigenvectors associated with each eigenvalue. Finding the eigenvectors is a simple matter of using Gaussian elimination to solve the equation

$$(\mathsf{A} - \lambda_i \, \mathrm{Id})\mathsf{v} = \vec{0}$$

for v, given the eigenvalue λ_i. It's important to note that with different eigenvalues we get different eigenvectors.[iii] So to get all the eigenvectors, we must consider all the eigenvalues, and solve this equation many times for differing λ_is. Note that there will not be a unique solution, since if v is an eigenvector, so is 2v, and πv, and any cv for a constant $c \neq 0$. To see this, note that

$$\mathsf{A}(c\mathsf{v}) = c\mathsf{A}\mathsf{v} = c\lambda\mathsf{v} = \lambda(c\mathsf{v})$$

for some eigenvalue λ.

Section 19. *Exercises*

V • Ex1. Exercise: Consider the matrix equation

$$\begin{bmatrix} 3 & -1 & -1 \\ -2 & 1 & -2 \\ 4 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -8 \\ -3 \\ 30 \end{bmatrix}$$

Use Gaussian elimination to solve for x, y, z.

V • Ex2. Exercise: Solve the matrix equation

$$\begin{bmatrix} -6 & 12 & 21 \\ -1 & 5 & 3 \\ 3 & -9 & -15 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 77 \\ 18 \\ -52 \end{bmatrix}$$

[iii] to see this, suppose v were an eigenvector associated to λ_1 and λ_2. Apply A, and get that $\mathsf{A}\mathsf{v} = \lambda_1 \mathsf{v} = \lambda_2 \mathsf{v}$, which means $(\lambda_1 - \lambda_2)\mathsf{v} = \vec{0}$. Since $\mathsf{v} \neq \vec{0}$, and $\lambda_1 - \lambda_2 \neq 0$, we get a contradiction, meaning there can be no such v.

V • Ex3. Exercise: Solve the matrix equation

$$\begin{bmatrix} -2 & 4 & -3 & 1 \\ -3 & 2 & 2 & 3 \\ 5 & 5 & 7 & -1 \\ -2 & -1 & 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} -42 \\ -19 \\ 33 \\ 13 \end{bmatrix}.$$

V • Ex4. Exercise: Find the inverse of the matrix

$$\begin{bmatrix} 4 & 3 \\ 2 & -6 \end{bmatrix}$$

V • Ex5. Exercise: Find the inverse of the matrix

$$\begin{bmatrix} -1 & 6 & 3 \\ 2 & 2 & 4 \\ 6 & -3 & 5 \end{bmatrix}$$

V • Ex6. Exercise: Solve the system of equations, or show that there is no solution.

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 1 \\ x_1 - x_2 + 2x_3 &= 1 \end{aligned}$$

V • Ex7. Exercise: Solve the system of equations, or show that there is no solution.

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 2 \\ 2x_1 + x_2 + x_3 &= 1 \\ x_1 - x_2 + 2x_3 &= -1 \end{aligned}$$

V • Ex8. Exercise: Determine whether the vectors

$$\mathsf{x}(t) = \langle e^{-t}, 2e^{-t} \rangle \qquad \mathsf{y}(t) = \langle e^{-t}, e^{-t} \rangle \qquad \mathsf{z}(t) = \langle 3e^{-t}, 0 \rangle.$$

are linearly independent for $-\infty < t < \infty$

V • Ex9. Exercise: Find all the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 5 & -1 \\ 3 & 1 \end{bmatrix}.$$

V • Ex10. Exercise: Find the eigenvalues and eigenvectors of the matrix

$$\mathsf{A} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}.$$

V • Ex11. Exercise: Suppose that $\mathsf{A}\mathsf{x} = 0$ for some vector $\mathsf{x} \neq 0$. Show there is some $\mathsf{y} \neq 0$ where $\mathsf{A}^*\mathsf{y} = 0$.

V • Ex12. Exercise: Solve the system of equations

$$\begin{aligned} x - 2y + z + u &= 1 \\ 2x + y - z - 2u &= 2 \\ 3x + 3y + 0z + 2u &= 3 \\ x + y + z + u &= 4 \end{aligned}$$

by Gaussian elimination.

V • Ex13. Exercise:

1. Solve by Gaussian elimination the system

$$\begin{aligned} x + 2y - z + u &= 0 \\ 2x - y + 3z + 2u &= 2 \\ x - 3y + z - u &= 3 \end{aligned}$$

$$-x + \quad y - \quad z + \quad u = a.$$

2. For what value of a does $u = 0$?

Chapter VI. Linear Systems of Differential Equations

We will look at differential equations which make statements about multiple functions. Most of the time, these functions will interact in some way, in the sense that the equalities depend on their values together.

In particular, we have the following definition for what we mean by a "system of differential equations".

19 • 1. Definition

A *system of differential equations* is a sequence of equations of the form

$$\begin{aligned} x_1' &= F_1(x_1, \cdots, x_n, t) \\ x_2' &= F_2(x_1, \cdots, x_n, t) \\ &\vdots \\ x_n' &= F_n(x_1, \cdots, x_n, t). \end{aligned}$$

A *solution* to such a system is a sequence of functions $x_1(t), x_2(t), \cdots, x_n(t)$ which satisfy those equations.

We will be interested in equations that are *linear* in form, as well as those which are *homogeneous*. These definitions are analogous to those in a single variable, and with the language of matrices, these definitions become practically the same. This is because a system of linear differential equations will be a sequence of equations

$$\begin{aligned} x_1' &= p_{11}(t)x_1 + \cdots + p_{1n}(t)x_n + g_1(t) \\ x_2' &= p_{21}(t)x_1 + \cdots + p_{2n}(t)x_n + g_2(t) \\ &\vdots \\ x_n' &= p_{n1}(t)x_1 + \cdots + p_{nn}(t)x_n + g_n(t). \end{aligned}$$

This can then be represented by the matrix equation

$$\begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} = \begin{bmatrix} p_{11}(t) & \cdots & p_{1n}(t) \\ \vdots & \ddots & \vdots \\ p_{n1}(t) & \cdots & p_{nn}(t) \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{bmatrix},$$

for various functions p_{ij} and g_i where $i, j \leq n$. In other words,

$$\mathsf{z}' = \mathsf{A}(t) \cdot \mathsf{z} + \mathsf{G}(t),$$

for matrices of functions $\mathsf{A}(t)$ and $\mathsf{G}(t)$. The derivative z' is just component-wise: we differentiate each entry in z.

19 • 2. Definition

A system of differential equations is *linear* iff it is of the form

$$\mathsf{z}' = \mathsf{A}(t) \cdot \mathsf{z} + \mathsf{G}(t),$$

for matrices of functions $\mathsf{A}(t)$ and $\mathsf{G}(t)$, and vector of variables z.

A system of differential equations is *homogeneous* iff $\mathsf{G}(t)$ is constantly 0 above: it's of the form

$$\mathsf{z}' = \mathsf{A}(t) \cdot \mathsf{z}.$$

So in the case that A is a constant matrix—i.e. just filled with constants—this situation simplifies considerably, as we'll see later.

For now, these equations can be useful for thinking about higher order differential equations, since we can go back and forth between the two for systems of linear equations.

19 • 3. Result

Every linear n-th order differential equation can be represented as a system of linear differential equations of n variables, and the reverse holds under certain conditions.

Proof .:.

For one direction, consider the linear n-th order differential equation

$$p_n(t)x^{(n)} + \cdots + p_1(t)x' + p_0(t)x = g(t).$$

The system of differential equations

$$\begin{aligned} x &= x_0 \\ x_0' &= x_1 \\ x_1' &= x_2 \\ &\vdots \\ x_{n-2}' &= x_{n-1} \\ p_n(t)x_{n-1}' + p_{n-1}(t)x_{n-1} \cdots + p_1(t)x_1 + p_0(t)x_0 &= g(t) \\ \longleftrightarrow x_{n-1}' &= \frac{1}{p_n(t)}(g(t) - p_{n-1}(t)x_{n-1} - \cdots - p_1(t)x_1 - p_0(t)x_0). \end{aligned}$$

then represents the n-th order differential equation, since x_i just represents the ith derivative of x.

The other direction is a lengthy process of differentiating equations, solving for a variable to reduce to fewer variables, and so on until we finish. So this will be a proof by example in the 2×2 case in Appendix A. ⊣

Section 20. *Homogeneous Systems with Constant Coefficients*

When dealing with systems of differential equations, a base case to consider is when the equations are linear, meaning of the form

$$\mathsf{z}' = \mathsf{A}(t) \cdot \mathsf{z} + \mathsf{G}(t)$$

for matrices A and G depending on t. To simplify this situation, we will consider for now equations where these are constant, and where $\mathsf{G} = 0$. In other words, the equation

$$\mathsf{z}' = \mathsf{A} \cdot \mathsf{z}$$

for some (constant) matrix A. To simplify the situation even further, we'll work just with 2×2 systems for now.

20 • 1. Result

Let A be a 2×2 matrix, and $\mathsf{z}(t)$ a vector. Consider the differential equation

$$\mathsf{z}' = \mathsf{A} \cdot \mathsf{z}.$$

Suppose A has distinct eigenvalues λ_1 and λ_2 with some associated eigenvectors v and w respectively. Thus solutions to the above differential equation are of the form

$$\mathsf{z}(t) = C\mathsf{v}e^{\lambda_1 t} + D\mathsf{w}e^{\lambda_2 t},$$

for some constants C and D.

Proof .:.

The proof of this can be found in Appendix A.

This generalizes to complex valued eigenvectors and eigenvalues, but often we want to work only with real-valued functions. In this case, we can simplify the answer using Euler's formula, and choosing different constants.

20 • 2. Result

Let A be a 2×2 matrix with real entries, and $\mathsf{z}(t)$ a vector. Consider the differential equation

$$\mathsf{z}' = \mathsf{A} \cdot \mathsf{z}.$$

Suppose A has complex eigenvalue $a + ib$ ($b \neq 0$) and associated eigenvector $\vec{v} + i\vec{w}$.

Thus the eigenvectors of A are $a \pm ib$ with eigenvectors $\vec{v} \pm i\vec{w}$, and solutions to the differential equation are
$$\mathsf{z}(t) = Ce^{at}(\vec{v}\cos(bt) - \vec{w}\sin(bt)) + De^{at}(\vec{v}\sin(bt) + \vec{w}\cos(bt)),$$
for some constants C and D.

Unfortunately, this is the simplest we can do. In the single variable case we get $Ce^{at}\sin(bt) + De^{at}\cos(bt)$, which corresponds to this under certain choices of C, D, and scaling $\vec{v}$ and $\vec{w}$.[i] Unfortunately, this just makes the other entry in z uglier. So to be fair, both are given an equal amount of ugliness. In practice, however, the ugliness will often be skewed towards one solution with the other being merely $Ce^{at}\sin(bt) + De^{at}\cos(bt)$.

Just like with a single variable, if there is only one eigenvalue—or solution to the characteristic equation in the case of a single variable—the solution isn't necessarily like in Result 20 • 1. And the result isn't so simply stated: the solution actually depends on whether the eigenvectors are linearly independent, and we get a somewhat different form from Result 20 • 1 if they are linearly dependent.

20 • 3. Result (2 x 2 Systems)

Let A be a 2×2 matrix, and $\mathsf{z}(t)$ a vector. Consider the differential equation
$$\mathsf{z}' = \mathsf{A}\mathsf{z}.$$
Suppose A has eigenvalue λ_1 with eigenvector $\mathsf{v}_1 \neq \vec{0}$, and eigenvalue λ_2 with eigenvector $\mathsf{v}_2 \neq \vec{0}$.

1. If $\lambda_1 \neq \lambda_2$, then
$$\mathsf{z}(t) = C\mathsf{v}_1 e^{\lambda_1 t} + D\mathsf{v}_2 e^{\lambda_2 t}.$$
2. If $\lambda_1 = \lambda_2 = \lambda$, but v_1 and v_2 are linearly independent, then
$$\mathsf{z}(t) = C\mathsf{v}_1 e^{\lambda t} + D\mathsf{v}_2 e^{\lambda t}.$$
3. If $\lambda_1 = \lambda_2 = \lambda$, and v_1 and v_2 are linearly *dependent*, then
$$\mathsf{z}(t) = C\mathsf{v}_1 e^{\lambda t} + D(\mathsf{v}_2 \cdot t \cdot e^{\lambda t} + \mathsf{m}e^{\lambda t}),$$
where m is a vector solving the matrix equation $(\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{m} = \mathsf{v}_2$.

Where C and D are constants.

Note that the choice of v_2 in the equation of m in case (3) was arbitrary. Because v_1 and v_2 were linearly dependent, they were a constant multiple of eachother: $\mathsf{v}_1 = k\mathsf{v}_2$ for some $k \neq 0$. In this case, we could just as well say that v is the only eigenvector (up to scaling), and write
$$\mathsf{z}(t) = C\mathsf{v}e^{\lambda t} + D(\mathsf{v} \cdot t \cdot e^{\lambda t} + \mathsf{m}e^{\lambda t})$$
for m solving $(\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{m} = \mathsf{v}$. This realy just results in absorbing the constant into D. The reason why such an m pops up is that we need a fundamental set of solutions, just like in the single variable case.

Now as a process of solving such systems, 2×2 Systems (20 • 3) translates to the series of steps:

1. Find the eigenvalues of A.
2. Find the corresponding eigenvectors of A.
3. Determine if the eigenvectors are linearly independent or linearly dependent.
4. If linearly dependent solve $(\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{m} = \mathsf{v}$ for an eigenvector v.
5. Use the appropriate solution as in 2×2 Systems (20 • 3).

More concretely, we get the series of steps on how to do the above.

1. Solve $\det(\mathsf{A} - \lambda\,\mathrm{Id}) = 0$ for λ by solving a polynomial equation.

2. For each λ above, solve $(\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{v} = \vec{0}$ for v using Gaussian elimination.
   - If $\lambda_1 \neq \lambda_2$, then $\mathsf{z}(t) = C\mathsf{v}_1 e^{\lambda_1 t} + D\mathsf{v}_2 e^{\lambda_2 t}$.
   - If $\lambda_1 = \lambda_2 = \lambda$, but v_1 and v_2 aren't multiples of each other, then $\mathsf{z}(t) = C\mathsf{v}_1 e^{\lambda t} + D\mathsf{v}_2 e^{\lambda t}$.

[i] In particular, if we fix some row, we're scaling $\vec{v} + i\vec{w}$ to make the entry in that row 1, and then dealing with the new eigenvectors where the complex values only occur in the other row.

3. If there is only one v up to scaling, then solve $(\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{m} = \mathsf{v}$ for m using Gaussian elimination.
   - In this case, $\mathsf{z}(t) = C\mathsf{v}e^{\lambda t} + D(\mathsf{v} \cdot t \cdot e^{\lambda t} + \mathsf{m}e^{\lambda t})$.

Note that having one eigenvalue with linearly independent eigenvectors is pretty rare. In fact, for the 2×2 case, this happens only when A is a multiple of the identity matrix.[ii]

Let's continue with 2×2 systems for a moment to deal with complex values.

20 • 4. Result (Complex 2 × 2 Solutions)

Let A be a 2×2 matrix with real entries, and $\mathsf{z}(t)$ a vector. Consider the differential equation

$$\mathsf{z}' = \mathsf{A} \cdot \mathsf{z}.$$

Suppose A has complex eigenvalue $a + ib$ ($b \neq 0$) and associated eigenvector $\mathsf{v} + i\mathsf{w}$.

Thus the eigenvectors of A are $a \pm ib$ with eigenvectors $\mathsf{v} \pm i\mathsf{w}$, and solutions to the differential equation are

$$\mathsf{z}(t) = Ce^{at}(\mathsf{v}\cos(bt) - \mathsf{w}\sin(bt)) + De^{at}(\mathsf{v}\sin(bt) + \mathsf{w}\cos(bt)),$$

for some constants C and D.

Showing all of these facts can be done just by converting them into homogeneous equations with constant coefficients, each of a single variable, and then solving them in the usual way we do there. In the case of complex eigenvalues, we merely use Euler's formula, and change the constants to remove any i.

Let's generalize this idea to $n \times n$ systems. Most larger systems will still be fairly small. Even the small 3×3 is still computationally difficult to deal with. But the ideas from the 2×2 systems generalize to $n \times n$ systems easily in the case that the eigenvalues are distinct.

20 • 5. Result ($n \times n$ systems)

Let A be an $n \times n$ matrix, and $\mathsf{z}(t)$ a vector. Consider the differential equation

$$\mathsf{z}' = \mathsf{A}\mathsf{z}.$$

Suppose A has eigenvalues λ_1, λ_2, ..., and λ_n—which aren't necessarily different—with corresponding eigenvectors $\mathsf{v}_1, \mathsf{v}_2, \cdots, \mathsf{v}_n$.

If the eigenvectors $\mathsf{v}_1, \cdots, \mathsf{v}_n$ are linearly independent *or* the eigenvalues $\lambda_1, \cdots, \lambda_n$ are all distinct, then

$$\mathsf{z}(t) = C_1\mathsf{v}_1e^{\lambda_1 t} + C_2\mathsf{v}_2e^{\lambda_2 t} + \cdots + C_n\mathsf{v}_ne^{\lambda_n t},$$

for constants $C_1, \cdots, C_n$.

If we have an eigenvalue λ repeated twice with only one (linearly independent) eigenvector, then instead of contributing the term $C\mathsf{v}e^{\lambda t} + D\mathsf{v}e^{\lambda t}$ to $\mathsf{z}(t)$, we add to $\mathsf{z}(t)$

$$C\mathsf{v}e^{\lambda t} + D(\mathsf{v}te^{\lambda t} + \mathsf{m}), \qquad \text{where} \qquad (\mathsf{A} - \lambda\,\mathrm{Id})^2\mathsf{m} = \vec{0} \neq \mathsf{m}.$$

If you go on to study more linear algebra, m is called a *generalized eigenvector* in that $(\mathsf{A} - \lambda\,\mathrm{Id})^k\mathsf{m} = \vec{0}$ for some k. In the case that $k = 1$, satisfying this just means being eigenvector associated to λ—so long as the vector isn't $\vec{0}$.

To continue this aside, things get more complicated as more and more eigenvalues are repeated. If an eigenvalue λ is repeated n times with one eigenvector v (up to scaling), then instead of adding $C_i\mathsf{v}e^{\lambda t}$ n times—yielding only one term when we need m—we add instead

$$C_ie^{\lambda t}\sum_{k=1}^{n}\left(\mathsf{m}_k\frac{t^k}{k!}\right),$$

for each $i \leq n$. Each m_k is defined by recursion:

$$\mathsf{m}_0 := \mathsf{v}_i$$
$$\mathsf{m}_1 := \text{a solution to } (\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{m} = \mathsf{m}_0$$

[ii]To see this, otherwise $\mathsf{A} - \lambda\,\mathrm{Id}$ isn't the $\mathbf{0}$ matrix. And thus we have an equation with at least one unknown: $(\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{v} = \vec{0}$. But then either one variable is fixed (e.g. if one equation only involves one unknown), or a and b have some sort of scalar relationship, which tells us that the eigenvectors can't be linearly independent.

$$\vdots$$
$$\mathsf{m}_{k+1} := \text{a solution to } (\mathsf{A} - \lambda\,\mathrm{Id})\mathsf{m} = \mathsf{m}_k.$$
$$\vdots$$

So if we're working in a 3×3 system, and we have eigenvalues λ_1 and $\lambda_2 = \lambda_3$, having eigenvectors v_1 and v_2 respectively, then the general solution will be

$$\mathsf{z}(t) = C\mathsf{v}_1 e^{\lambda_1 t} + De^{\lambda_2 t}(t\mathsf{v}_2 + \mathsf{m}).$$

If we have just one eigenvalue $\lambda = \lambda_1 + \lambda_2 + \lambda_3$ with just one eigenvector v, then the solution will be

$$\mathsf{z}(t) = C_1\mathsf{v}e^{\lambda t} + C_2\mathsf{v}e^{\lambda t}(t\mathsf{v} + \mathsf{m}_1) + C_3 e^{\lambda t}\left(\frac{t^2}{2} + t\mathsf{m}_1 + \mathsf{m}_2\right).$$

In general, however, it's simpler to refer to notions of "exponentiation" with matrices defined along the lines of a power series used in the case of real numbers. In this case, $\mathsf{z}' = \mathsf{A}\mathsf{z}$ has a solution of $\exp(\mathsf{A}t)\vec{c}$ for a (column) vector of constants $\vec{c}$.

Section 21. *Non-Homogeneous Systems*

Many of the results about linear, differential equations—of a single variable—apply also to *systems* of such equations. In particular, the general solution of a homogeneous $n \times n$ system is one where we have n "different enough" solutions ϕ_1, through ϕ_n, in that they have a non-zero Wronskian matrix. The general solution in that case is just $C_1\phi_1 + \cdots + C_n\phi_n$. If we have a non-homogeneous system, the difference is still a single, particular solution:

21 • 1. Result

Consider the $n \times n$ linear system of differential equations

$$\mathsf{z}' = \mathsf{A}(t) \cdot \mathsf{z} + \mathsf{G}(t).$$

Let $\mathsf{Z}(t)$ be any solution to this. Let $\phi_1, \phi_2, \cdots, \phi_n$ solve the homogeneous version: $\phi' = \mathsf{A}(t) \cdot \phi$.

Thus the general solution is

$$\mathsf{z}(t) = C_1\phi_1(t) + C_2\phi_2(t) + \cdots + C_n\phi_n(t) + \mathsf{Z}(t)$$

for constants C_1, ..., C_n.

Proof .:.

Consider an arbitrary solution z of the non-homogeneous differential equation. Consider the function $\Phi = \mathsf{z} - \mathsf{Z}$. Note that then

$$\begin{aligned}\Phi' &= \mathsf{z}' - \mathsf{Z}' \\ &= \mathsf{A}(t) \cdot \mathsf{z} + \mathsf{G}(t) - (\mathsf{A}(t) \cdot \mathsf{Z} + \mathsf{G}(t)) \\ &= \mathsf{A}(t) \cdot (\mathsf{z} - \mathsf{Z}) \\ &= \mathsf{A}(t) \cdot \Phi.\end{aligned}$$

Hence Φ is a solution to the homogeneous version. Hence $\Phi = \mathsf{z} - \mathsf{Z}$ is of the form $\Phi = C_1\phi_1 + \cdots + C_n\phi_n$ for constants C_1, ..., C_n. But then, adding Z to both sides, $\mathsf{z} = C_1\phi_1 + \cdots + C_n\phi_n + \mathsf{Z}$. ⊣

So to solve non-homogeneous equations, we have similar methods as in the single variable case. In particular, what follows is the $n \times n$ version of undetermined coefficients.

21 • 2. Result

Consider the $n \times n$ linear system of differential equations $\mathsf{z}' = \mathsf{A} \cdot \mathsf{z} + \mathsf{G}(t)$.

If $g(t) = \vec{c}_n t^n + \cdots + \vec{c}_1 t + \vec{c}$,	then guess	$Y(t) = \vec{C}_n t^n + \cdots + \vec{C}_1 t + \vec{C}_0$;
if $g(t) = \vec{c}\sin(at)$, or $g(t) = \vec{c}\cos(at)$,	then guess	$Y(t) = \vec{C}\sin(at) + \vec{D}\cos(at)$;
if $g(t) = \vec{c}e^{at}$,	then guess	$Y(t) = \vec{C}e^{at}$;

If this guess doesn't work—i.e. it's a solution to the homogeneous version—then keep multiplying the guess by t—including lower degrees—until it doesn't.

For example, you might initially guess $\vec{c}e^{at}$. If that doesn't work, you would consider not just $\vec{c}te^{at}$, but $\vec{c}te^{at} + \vec{d}e^{at}$. If that doesn't work, you would then consider $\vec{c}_1 t^2 e^{at} + \vec{c}_2 te^{at} + \vec{c}_3 e^{at}$. And so on and so forth until it does work.

Note also that, just like the single variable case, undetermined coefficients can only be used in the case that the linear system has constant coefficients.

Section 22. *Linearization*

There are many more kinds of differential equations beyond just ones of the form
$$\mathsf{z}' = \mathsf{A}(t)\mathsf{z} + \mathsf{G}(t),$$
for matrix functions A and G. We really only know, in principle, how to solve them when A is constant, and G is nice enough. But more generally, systems of differential equations are of the form
$$\mathsf{z}' = F(t, \mathsf{z}),$$
with no real restrictions on F.[iii] Let's consider the differential equation of this form
$$\mathsf{z}' = \begin{bmatrix} F_1(\mathsf{z}) \\ F_2(\mathsf{z}) \\ \vdots \\ F_n(\mathsf{z}) \end{bmatrix},$$
where we at least have that each F_i is differentiable and doesn't depend on t. We can *approximate* such a solution using an approximation to the differential equation. To do this, assume $F(\mathsf{s}) = 0$ at some (vector) value s. Now we use a linear approximation to F at s, and consider the new, linear system of differential equations around s.

That's a lot to take in, so let's just consider the 2×2 case:
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} f(x, y) \\ g(x, y) \end{bmatrix}$$
for differentiable functions f and g. Suppose $f(x_1, y_1) = g(x_1, y_1) = 0$ for numbers x_1 and y_1. Such a point $\langle x_1, y_1\rangle$ is a *critical point* for $\langle f, g\rangle$. Because of this, we have approximations to $f(x, y)$ and $g(x, y)$ near $\mathsf{s} = \langle x_1, y_1\rangle$ by a *Taylor approximation*: denoting $\partial f/\partial x$ by f_x and so forth,
$$\begin{aligned} f(x, y) &\approx f(\mathsf{s}) + f_x(\mathsf{s})(x - x_1) + f_y(\mathsf{s})(y - y_1) \\ &\approx f_x(\mathsf{s})(x - x_1) + f_y(\mathsf{s})(y - y_1), \\ g(x, y) &\approx g(\mathsf{s}) + g_x(\mathsf{s})(x - x_1) + g_y(\mathsf{s})(y - y_1) \\ &\approx g_x(\mathsf{s})(x - x_1) + g_y(\mathsf{s})(y - y_1). \end{aligned}$$
Hence we can write a different system of equations whose solution is *not* the same as before, but which is supposed to approximate the solution to $\mathsf{z}' = F(\mathsf{z})$ when z is near s: choose new variables $u = x - x_1$, $v = y - y_1$, and consider the new system
$$\begin{bmatrix} u' \\ v' \end{bmatrix} \approx \begin{bmatrix} f_x(\mathsf{s}) & f_y(\mathsf{s}) \\ g_x(\mathsf{s}) & g_y(\mathsf{s}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$
the system above—replacing $\approx$ with $=$—is called the *linearization* of the original system. The variables u and v are

[iii] beyond the fact that it outputs a vector the size of z.

approximations to $x-x_1$ and $y-y_1$ near 0. The matrix on the left is the *Jacobian* of F, denoted $\mathsf{J}(F)$, evaluated at s.

22 • 1. Definition

Consider the differential equation

$$\mathsf{z}' = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = F(\mathsf{z})$$

for continuously differentiable functions f and g. Suppose $F(\mathsf{s}) = \vec{0}$ for vector $\mathsf{s} = \langle x_1, y_1\rangle$, a *critical point* of F.

The *linearization* of $\mathsf{z}' = F(\mathsf{z})$ is the linear system

$$\mathsf{w}' = \mathsf{J}(F)(\mathsf{s}) \cdot \mathsf{w} = \begin{bmatrix} f_x(x_1,y_1) & f_y(x_1,y_1) \\ g_x(x_1,y_1) & g_y(x_1,y_1) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix},$$

where $\langle u, v\rangle = \mathsf{w} \approx \mathsf{z}-\mathsf{s} = \langle x-x_1, y-y_1\rangle$.

Just like with autonomous differential equations—equations of the form $y' = f(y)$—there is a notion of stability for critical points of $\mathsf{z}' = F(\mathsf{z})$. This notion is slightly complicated by the fact that we have many functions to worry about, but ultimately, the concept is similar. Consider first the simple case of when we're dealing with homogeneous equations.

So long as $\det(\mathsf{A}) \neq 0$, the only critical point of $\mathsf{z}' = \mathsf{A}\mathsf{z}$ is $\mathsf{z} = \vec{0}$. This point is *asymptotically stable* if solutions z go to this critical point $\vec{0}$ as $t \to \infty$. We know that solutions to such equations are usually of the form

$$C_1 \mathsf{v}_1 e^{\lambda_1 t} + \cdots + C_n \mathsf{v}_n e^{\lambda_n t}.$$

Given that all those $C_i \mathsf{v}_i$ are all just constant vectors, the real determiner of where the solutions go is all those $e^{\lambda_i t}$s, and so the eigenvalues. If all the eigenvalues have a real part that's negative, then such solutions go to $\vec{0}$.

So the notion of stability is roughly the same, and it is contingent on whether functions that make up the solution go to 0, or go to ∞. This is really determined by the eigenvalues of the matrix $\mathsf{J}(F)(\mathsf{s})$.

22 • 2. Definition

Consider the differential equation $\mathsf{z}' = \mathsf{A}\mathsf{z}$. Let λ_1, and λ_2 be eigenvalues for A.

1. The critical point $\vec{0}$ is *asymptotically stable* iff $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ both go to 0 as $t \to \infty$. In other words, λ_1 and λ_2 (or at least their real part) are negative.
2. The critical point $\vec{0}$ is *stable* iff $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ are each constant or have no limit (not ∞, and not 0) as $t \to \infty$. In other words, λ_1 and λ_2 (or at least their real part) are 0.
3. The critical point $\vec{0}$ is *unstable* iff $e^{\lambda_1 t}$ or $e^{\lambda_2 t}$ goes to ∞ as $t \to \infty$. In other words, λ_1 or λ_2 (or at least their real part) is positive.

If we look back at Definition 22 • 1, if $\mathsf{w} = \vec{0}$ is an asymptotically stable critical point of the linearization, then $\mathsf{w} \approx \mathsf{z}-\mathsf{s}$ goes to $\vec{0}$ as $t \to \infty$, i.e. z converges to s. Similarly, if w goes to ∞, then $\mathsf{z}-\mathsf{s}$ does not go to 0, and so z diverges from s. This is just to say that by looking at the linearization, we can tell whether solutions near z converge to s or not.

22 • 3. Result

Consider the differential equation $\mathsf{z}' = F(\mathsf{z})$ with linearization $\mathsf{w}' = \mathsf{J}(F)(\mathsf{s}) \cdot \mathsf{w}$ around the critical point s. Therefore,

1. the critical point $\mathsf{z} = \mathsf{s}$ is asympotically stable if $\mathsf{w} = \vec{0}$ is; and
2. the critical point $\mathsf{z} = \mathsf{s}$ is unstable if $\mathsf{w} = \vec{0}$ is.

Note that this is really all determined by our choice of s. We look at the linearization $(\mathsf{z}-\mathsf{s})' \approx \mathsf{A}(\mathsf{z}-\mathsf{s})$, and see whether $\mathsf{z}-\mathsf{s} = \vec{0}$ is an asymptotically stable or unstable critical point, and thus whether solutions z near s go to s or not.

Section 23. Exercises

VI • Ex1. Exercise: Transform the given initial value problem into an initial value problem for two first-order equations.

$$u'' + p(t)u' + q(t)u = g(t), u(0) = u_0, u'(0) = u'_0.$$

VI • Ex2. Exercise: Consider the system

$$\begin{aligned} x'_1 &= -2x_1 + x_2 \\ x'_2 &= x_1 - 2x_2. \end{aligned}$$

a. Solve the first differential equation for x_2.
b. Substitute the result of (a) into the second differential equation, thereby obtaining a second-order differential equation for x_1.
c. Solve the differential equation in (b) for x_1.
d. Use (a) and (c) to find x_2.

§ 23 A. 2 × 2 Systems

VI • Ex3. Exercise: Solve the differential equation

$$\mathsf{z}' = \mathsf{A}\mathsf{z} = \begin{bmatrix} 3 & 2 \\ -1 & 6 \end{bmatrix} \mathsf{z}$$

for the vector function $\mathsf{z}(t)$.

VI • Ex4. Exercise: Solve the differential equation

$$\mathsf{z}' = \mathsf{A}\mathsf{z} = \begin{bmatrix} i+1 & 0 \\ -2 & 3 \end{bmatrix} \mathsf{z}$$

for the vector function $\mathsf{z}(t)$.

VI • Ex5. Exercise: Solve the differential equation

$$\mathsf{z}' = \mathsf{A}\mathsf{z} = \begin{bmatrix} 8 & 4 \\ -4 & 6 \end{bmatrix} \mathsf{z}$$

for the vector function $\mathsf{z}(t)$.

VI • Ex6. Exercise (Two Eigenvalues): Solve the system of differential equations

$$\begin{aligned} x' &= 3x + 2y \\ y' &= 2x + 3y, \end{aligned}$$

with initial conditions $x(0) = 2$, and $y(0) = 12$.

VI • Ex7. Exercise (One Eigenvalue, Two Eigenvectors): Solve the system of differential equations

$$\begin{aligned} x' &= 2x \\ y' &= 2y, \end{aligned}$$

with initial conditions $x(0) = 2$, and $y(0) = 1$.

VI • Ex8. Exercise (One Eigenvalue, One Eigenvector): Solve the system of differential equations

$$\begin{aligned} x' &= x + 4y \\ y' &= -4x + 9y, \end{aligned}$$

with initial conditions $x(0) = -1$, and $y(0) = 0$.

VI • Ex9. Exercise (Complex Eigenvalues): Solve the initial value problem

$$\mathsf{z}' = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \mathsf{A}\mathsf{z},$$

where $x(0) = -4$, and $y(0) = 1$.

VI • Ex10. Exercise: Solve the system of equations

$$\begin{aligned} x' &= 3x - y \\ y' &= x + 2y, \end{aligned}$$

where $x(0) = 0$, and $y(0) = 1$.

VI • Ex11. Exercise: Solve the system of equations

$$\begin{aligned} x' &= x - 2y \\ y' &= 2x + y, \end{aligned}$$

with initial conditions $x(0) = 1$, and $y(0) = 1$.

VI • Ex12. Exercise: Solve the non-homogeneous system of differential equations

$$\begin{aligned} x' &= x - y + \sin(t) \\ y' &= y + e^{2t} \end{aligned}$$

VI • Ex13. Exercise: Find the general solution to the system

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x + y + t \\ x + 2y \end{bmatrix}.$$

VI • Ex14. Exercise: Consider the non-linear system of differential equations

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x + \cos(y) \\ x \cdot y \end{bmatrix} = \begin{bmatrix} f(x, y) \\ g(x, y) \end{bmatrix}.$$

1. Find the critical points of the system;
2. give the linearization of this around the point $\mathsf{s} = \langle 0, \pi/2 \rangle$;
3. classify the stability that s has.

VI • Ex15. Exercise: Consider the following non-linear system

$$\begin{aligned} x' &= 3x - y + 3y^2 \\ y' &= x + 2y \end{aligned}$$

with initial conditions $x(0) = 1$, and $y(0) = 1$.

1. Find the critical points of the system.
2. Find the linear approximation around each point.
3. One critical point is $\langle 0, 0 \rangle$. What is the type of this critical point?
   NOTE: you only need to find the eigenvalues to answer this part.

§ 23 B. 3 × 3 Systems

Now let's consider some 3×3 systems, which are computationally more involved.

VI • Ex16. Exercise (Three Eigenvalues): Solve the initial value problem

$$\mathsf{z}' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 14 & 12 \\ -1 & 10 & 6 \\ 1 & -11 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathsf{A}\mathsf{z},$$

where $\mathsf{z}(0) = \langle 0, 3, 1 \rangle$.

VI • Ex17. Exercise (Two Eigenvalues in a 3 × 3 System): Solve the initial value problem

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -1 & 2 \\ 0 & 16 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

where $z(0) = \langle 200, 0, 100 \rangle$.

Appendix A. Unnecessary Proofs

A • 1. Result

Consider the system of linear differential equations $\mathsf{z}' = \mathsf{A}(t)\cdot\mathsf{z} + \mathsf{G}(t)$ written as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a(t) & b(t) \\ c(t) & d(t) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} g_1(t) \\ g_2(t) \end{bmatrix}.$$

If $b(t) \neq 0$ and $c(t) \neq 0$, we can write

$$x'' - \left(a + d + \frac{b'}{b}\right)x' + \left(ad - bc - a' - \frac{b'}{b}\right)x = g_1' - \left(\frac{b'}{b} + d\right)g_1 + bg_2$$
$$y'' - \left(d + a + \frac{c'}{c}\right)y' + \left(da - cb - d' - \frac{c'}{c}\right)y = g_2' - \left(\frac{c'}{c} + a\right)g_2 + cg_1.$$

In the case that a, b, c, d are all constants, we have

$$x'' - \mathrm{trace}(\mathsf{A})x' + \det(\mathsf{A})x = g_1' - dg_1 + bg_2$$
$$y'' - \mathrm{trace}(\mathsf{A})y' + \det(\mathsf{A})y = g_2' - ag_2 + cg_1.$$

Proof .:.

Consider the system of linear differential equations $\mathsf{z} = \mathsf{A}(t)\mathsf{z} + \mathsf{G}(t)$, also written

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a(t) & b(t) \\ c(t) & d(t) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} g_1(t) \\ g_2(t) \end{bmatrix}.$$

Understanding that a, b, c, d are all functions of t, I will now abandon writing "(t)" for the sake of space. Now the above equality means, so long as we can divide by b and c,

$$y = \frac{x' - ax - g_1}{b}, \qquad x = \frac{y' - dy - g_2}{c}.$$

differentiation on the matrix equation yields

$$\mathsf{z}'' = \mathsf{A}'\mathsf{z} + \mathsf{A}\mathsf{z}' + \mathsf{G}',$$

which can be written

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} g_1' \\ g_2' \end{bmatrix}.$$

substituting the equalities of y' in terms of y, then y in terms of x and x' above, this gives the equations

$$\begin{aligned} x'' &= a'x + ax' + g_1' + b'y + by' \\ &= a'x + ax' + g_1' + b'y + b(cx + dy + g_2) \\ &= \left(a' + \frac{b'}{b} + cb - ad\right)x + \left(\frac{b'}{b} + a + d\right)x' + g_1' - \left(\frac{b'}{b} + d\right)g_1 + bg_2, \end{aligned}$$

and we get a similar result for y:

$$y'' = \left(d' + \frac{c'}{c} + bc - da\right)y + \left(\frac{c'}{c} + d + a\right)y' + g_2' - \left(\frac{c'}{c} + a\right)g_2 + cg_1.$$

In the case that a, b, c, d are all constants, $a' = b' = c' = d' = 0$ and so this simplifies to the equations

$$x'' - \mathrm{trace}(\mathsf{A})x' + \det(\mathsf{A})x = g_1' - dg_1 + bg_2$$
$$y'' - \mathrm{trace}(\mathsf{A})y' + \det(\mathsf{A})y = g_2' - ag_2 + cg_1 \qquad \dashv$$

A • 2. Result

Let A be a 2×2 matrix, and $\mathsf{z}(t)$ a vector. Consider the differential equation

$$\mathsf{z}' = \mathsf{A}\cdot\mathsf{z}.$$

Suppose A has distinct eigenvalues λ_1 and λ_2 with some associated eigenvectors v and w respectively. Thus solutions to the above differential equation are of the form

$$\mathsf{z}(t) = C\mathsf{v}e^{\lambda_1 t} + D\mathsf{w}e^{\lambda_2 t},$$

for some constants C and D.

Proof .:.

First of all, anything of that form $C\mathsf{v}e^{\lambda_1 t} + D\mathsf{w}e^{\lambda_2 t}$ will be a solution to the differential equation. This is because

$$\begin{aligned}\mathsf{z}(t) = C\mathsf{v}e^{\lambda_1 t} + D\mathsf{w}e^{\lambda_2 t} \longrightarrow \mathsf{z}'(t) &= C\lambda_1\mathsf{v}e^{\lambda_1 t} + D\lambda_2\mathsf{w}e^{\lambda_2 t} \\ &= C\mathsf{A}\cdot\mathsf{v}e^{\lambda_1 t} + D\mathsf{A}\cdot\mathsf{w}e^{\lambda_2 t} \\ &= \mathsf{A}\cdot(C\mathsf{v}e^{\lambda_1 t} + D\mathsf{w}e^{\lambda_2 t}) \\ &= \mathsf{A}\cdot\mathsf{z}(t).\end{aligned}$$

So now we need to show that *any* solution to $\mathsf{z}' = \mathsf{A}\mathsf{z}$ is of this form. Write out the matrix A and vector v as

$$\mathsf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \mathsf{v} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Thus we have

$$x' = ax + by, \qquad \text{and} \qquad y' = cx + dy.$$

Now we consider two cases:

(1) $b = c = 0$. In this case, the differential equations have no interaction: $x' = ax$ and $y' = dy$ so that the solution is

$$\mathsf{z}(t) = \begin{bmatrix} Ce^{at} \\ De^{dt} \end{bmatrix},$$

for some constants C and D. Note that this agrees with the predicted form of the result. To see this, the eigenvalues of A are a and d. The eigen vectors are just $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ times any constant, and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ times any constant. Hence the predicted solution is

$$C\begin{bmatrix} c_1 \\ 0 \end{bmatrix}e^{at} + D\begin{bmatrix} 0 \\ c_2 \end{bmatrix}e^{dt} = \begin{bmatrix} C'e^{at} \\ D'e^{dt} \end{bmatrix}$$

for some constants C, C', D, D', c_1, and c_2. And this is precisely the form it must be in above.

(2) $b \neq 0$ or $c \neq 0$. For definiteness, say $b \neq 0$, since it's the same argument. This allows us to solve the first equation for y:

$$\begin{aligned} y &= \frac{1}{b}(x' - ax) \\ \therefore\ y' &= \frac{1}{b}x'' - \frac{a}{b}x'. \end{aligned}$$

The second equation tells us that $y' = cx + dy$ so that equating the two yields

$$\begin{aligned} cx + dy = \frac{1}{b}x'' - \frac{a}{b}x' &\longleftrightarrow cx + \frac{d}{b}(x' - ax) = \frac{1}{b}x'' - \frac{a}{b}x' \\ &\longleftrightarrow cbx + dx' - adx = x'' - ax' \\ &\longleftrightarrow x'' - \text{trace}(\mathsf{A})x' + \det(\mathsf{A})x = 0. \end{aligned}$$

This is a homogeneous, differential equation of a single variable. Solving this requires us to find r where

$$r^2 - \text{trace}(\mathsf{A})r + \det(\mathsf{A}) = 0.$$

Note, however, that λ is an eigenvalue iff λ satisfies

$$\begin{aligned} \det(\mathsf{A} - \lambda\,\mathrm{Id}) = 0 &\longleftrightarrow \det\left(\begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix}\right) = 0 \\ &\longleftrightarrow (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + ad - bc = 0 \\ &\longleftrightarrow \lambda^2 - \text{trace}(\mathsf{A})\lambda + \det(\mathsf{A}) = 0. \end{aligned}$$

Hence the r we find will be eigenvalues. Since we assumed there were two distinct eigenvalues λ_1 and λ_2, this gives

$$
\begin{aligned}
x(t) &= Ce^{\lambda_1 t} + De^{\lambda_2 t} \\
x'(t) &= \lambda_1 Ce^{\lambda_1 t} + \lambda_2 De^{\lambda_2 t} \\
y(t) &= \frac{1}{b}(x'(t) - ax(t)) \\
&= \frac{\lambda_1}{b}Ce^{\lambda_1 t} + \frac{\lambda_2}{b}De^{\lambda_2 t} - \frac{a}{b}Ce^{\lambda_1 t} - \frac{a}{b}De^{\lambda_2 t} \\
&= \left(\frac{\lambda_1 - a}{b}\right)Ce^{\lambda_1 t} + \left(\frac{\lambda_2 - a}{b}\right)De^{\lambda_2 t}. \\
\therefore \mathsf{z}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} &= C\begin{bmatrix} 1 \\ (\lambda_1 - a)/b \end{bmatrix} e^{\lambda_1 t} + D\begin{bmatrix} 1 \\ (\lambda_2 - a)/b \end{bmatrix} e^{\lambda_2 t}.
\end{aligned}
$$

So it suffices to show that

$$
\begin{bmatrix} 1 \\ (\lambda_1 - a)/b \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 1 \\ (\lambda_2 - a)/b \end{bmatrix}
$$

are eigenvectors with eigenvalues λ_1 and λ_2 respectively. To show that this is the case, we merely multiply these vectors by A on the left. Given that $\lambda^2 - (a + d)\lambda + (ad - bc) = 0$ for each λ, we get that $d\lambda - (ad - bc) = \lambda^2 - a\lambda$:

$$
\begin{aligned}
\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 \\ (\lambda - a)/b \end{bmatrix} = \begin{bmatrix} a + (\lambda - a) \\ c + (d\lambda - da)/b \end{bmatrix} &= \begin{bmatrix} \lambda \\ (d\lambda + cb - da)/b \end{bmatrix} \\
= \begin{bmatrix} \lambda \\ (\lambda^2 - a\lambda)/b \end{bmatrix} &= \lambda \begin{bmatrix} 1 \\ (\lambda - a)/b \end{bmatrix}
\end{aligned}
$$

Thus each vector is an eigenvector with the corresponding eigenvalue. This means that

$$
\begin{aligned}
\mathsf{z}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} &= C\begin{bmatrix} 1 \\ (\lambda_1 - a)/b \end{bmatrix} e^{\lambda_1 t} + D\begin{bmatrix} 1 \\ (\lambda_2 - a)/b \end{bmatrix} e^{\lambda_2 t} \\
&= C\mathsf{v}e^{\lambda_1 t} + D\mathsf{w}e^{\lambda_2 t}
\end{aligned}
$$

where v and w are eigenvectors with eigenvalues λ_1 and λ_2 respectively.

Reference

[1] William E. Boyce, Richard C. Diprima, and Douglas B. Meade, *Elementary Differential Equations and Boundary Value Problems*, 11th ed., John Wiley & Sons, Inc., 2017.

Index