MAT267: Advanced Ordinary Differential Equations

Rishibh Prakash

January 2022

Contents

1 Techniques for solving ODES

1.1 Separation of variables

Suppose we have a differential equation of the following form

$$\frac{dx}{dt} = f(t)g(x)$$

where f, g are continuous. Suppose we are also given that $x(0) = x_0$ as our initial condition.

If $g(x_0) = 0$ then $x(t) \equiv x_0$ is a solution. So suppose that $g(x_0) \neq 0$, implying that g is non-zero in some neighbourhood around x_0. We can thus do the following

$$\begin{aligned}
\frac{dx}{dt} &= f(t)g(x) \\
\frac{x'(t)}{g(x(t))} &= f(t) \\
\int \frac{x'(t)}{g(x(t))} dt &= \int f(t) dt \\
\int \frac{1}{g(u)} &= \int f(t) dt \\
G(x) &= F(t) + c
\end{aligned}$$

where in the penultimate line we substitute $u = x(t)$ and in the final line we take G to be the anti-derivative of $\frac{1}{g}$ and F to be the anti-derivative of f. Recall we are solving around $(0, x_0)$ and we know that g is non-zero in non-zero in a neighbourhood of it. Thus $G'(x) = \frac{1}{g(x)}$ is also non-zero (either all positive or all non-negative) implying that G is invertible on this neighbourhood. We can thus take $x = G^{-1}(F(t) + c)$. The solution passing through (t_0, x_0) corresponds to $c = G(x_0) - F(t_0)$.

Remark 1.1. The concrete construction of the solution implies that the solution to a differential equation with separable values is unique in a small neighbourhood of (t_0, x_0), given than $g(x_0) \neq 0$.

The above steps are often written more simply as

$$\begin{aligned}
\frac{dx}{dt} &= f(t)g(x) \\
\frac{1}{g(x)} dx &= f(t) dt \\
\int \frac{1}{g(x)} dx &= \int f(t) dt \\
G(x) = F(t) + c
\end{aligned}$$

1.1.1 Example

A simple example of using separation of variables is Equation 3.3 (no more guesswork required!). A more interesting example can be found in Section 5.

1.2 Homogeneous Functions

We say a function F on $\mathbb{R}^2$ is homogeneous of degree α (with $\alpha \in \mathbb{R}$) if $F(tx, ty) = t^\alpha F(x, y)$.

Suppose we are given

$$P(x, y)dx + Q(x, y)dy = 0 \tag{1.1}$$

Where P and Q are homogeneous and of the same degree. Then substituting $y = xu$ (giving us $dy = udx + xdu$) or $x = yv$ (giving us $dx = vdy + ydv$) changes our differential equation into one where we can separate variables.

1.2.1 Example

Suppose we have the following equation

$$\underbrace{\left[xe^{\frac{y}{x}} - y\sin\left(\frac{y}{x}\right)\right]}_{P(x,y)} dx + \underbrace{x\sin\left(\frac{y}{x}\right)}_{Q(x,y)} dy = 0 \tag{1.2}$$

It is easy to verify P and Q are homogeneous functions (of degree 2). We substitute $y = xu$ (either substitution works but often one is easier that the other. In this case it seems quite apparent that one would want to replace $\frac{y}{x}$). The substitution gives us

$$\begin{aligned} [xe^u - xu\sin(u)]dx + x\sin(u)(udx + xdu) = 0 \\ e^u dx + x\sin(u)du = 0 \end{aligned}$$

(note we divide by x without worry since the original equation Equation 1.2 doesn't allow for $x = 0$) The variables can now clearly be separated.

2 Introduction

Differential equations are equations where the unknown is a function and the given equation gives us a relationship between the function and its derivative(s). Such equations are common in physics, economomics, etc. where we know how a quantity changes (i.e. its derivatives) but not necessarily how to determine the quantity at any given instant. Perhaps the most important differential equation is Newton's (second) law:

$$ma = F$$

This equation tells us how force and acceleration, the second derivative of position, are interrelated. A similar example is that of the spring, where Hooke's law (in combination with Newton's law) tells us:

$$m \cdot x''(t) = -k \cdot x(t) \tag{2.1}$$

Without too much effort, we see that

$$x(t) = \cos\left(\sqrt{\frac{k}{m}}t\right)$$

forms a solution to (2.1), by which we mean that it satisfies the differential equation. Playing around a bit more we find that

$$x(t) = A\cos\left(\sqrt{\frac{k}{m}}t\right) + B\sin\left(\sqrt{\frac{k}{m}}t\right) \tag{2.2}$$

where A and B are some real numbers, all form solutions to (2.1). As it turns out, this is what *all* solutions to (2.1) look like, although this is something we will prove later.

The parameters A and B are determined by the initial conditions that the differential equation needs to satisfy (in general 2 unknowns will require 2 initial conditions). Hence A and B in some sense parametrise the solution space. In such cases, we would like to have the parameters cover the entire solution space and for each set of parameters to correspond to a different solution.

2.1 Solving Differential Equations

There are two types of differential equations: ordinary differential equations, ODEs, (where the unknown function is in one variable) and partial differential equations, PDEs, (where the unknown function is in several variables). Our goal is to understand the quantity that the unknown function measures. Of course the best way to understand this quantity is by finding the unknown function. For example, the solutions above (see Equation 2.2) instantly tell us that x has periodic behaviour, not something immediate from the differential equation itself.

This is why there is so much interest in solving differential equations. ODEs can sometimes be solved analytically (see the above example), PDEs, almost never. However, we can often analyse differential equations themselves in order to make qualitative statements about the functions (how it changes, its limiting behaviour, points of equilibrium, etc.) and still learn meaningful information about the quantity being measured.

But perhaps we go any further, we should probably define what ODEs are

Definition 2.1 (Ordinary Differential Equation)**.** An ODE is an equation of the form

$$F(t, x(t), \dots, x^{(k)}(t)) = 0$$

where x is a vector valued function on an open interval $I \subset \mathbb{R}$ which is $k-$times continuous differentiable[1]. This is known as the implicit form of the ODE.

If the codomain of F is $\mathbb{R}^m$ with $m > 1$, we get a system of equations. Sometimes we can express the $k-$th derivative as a function of the lower order derivatives which gets the standard or explicit form.

We should also probably define what it means to solve an ODE.

Definition 2.2. A (classical) solution of an ODE $F(t, x(t), \dots, x^{(k)})(t) = 0$ is a function of class C^k $\phi : I \to \mathbb{R}^n$, where $I \subset \mathbb{R}$ is an open interval, such that $F(t, \phi(t), \dots, \phi^{(k)}(t)) = 0$ for all $t \in I$.

Note that not all ODEs can be solved. Consider, for example,

$$\left|\frac{dy}{dx}\right| + |y| + 1 = 0$$

Now, consider the following non-example. Suppose we are given the ODE

$$x + y \cdot y' = 0$$

We can define $y(x) = \sqrt{-(1+x^2)}$ and we see that

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

which means that $x + y \cdot y'$ is certainly equal to 0. However, y is not defined in the reals!

Definition 2.3. The general solution of an ODE is a formula for all possible solutions.

The solution for x above, (2.2), is an example of a general solution. It is normally no easy task to find the general solution to a differential equation as one needs to prove that they have indeed found *all* possible solutions.

2.2 Standard Trick

There is a standard trick to turn a higher order ODE into a system of first-order ODEs which is a bit simple-minded but occassionally useful.

Example 2.1. Suppose we now have the equation

$$mx'' = -kx - cx'$$

(one may think of this as introducing drag into our spring example). We can introduce new variables

$$x_1 = x, x_2 = x'$$

allowing us to write

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} x_2 \\ -\frac{-k}{m}x_1 - \frac{c}{m}x_2 \end{pmatrix}$$

■

[1] In principle we don't need continuity of the $k-$th derivative, only its existence. However this does make things a lot nicer

In the general case of $F(t, x, x', \dots, x^{(k)}) = 0$, we define

$$\begin{aligned} x_1 &= x \\ x_2 &= x_1' \\ x_3 &= x_2' \\ &\vdots \\ x_k &= x_{k-1}' \end{aligned}$$

thus allowing us to write $F(t, x_1, x_2, \dots, x_k) = 0$.

3 Simple Examples

We list a few 'classic' examples of ODEs here.

The first example is perhaps the simplest one, one could think of

$$x' = 0 \tag{3.1}$$

It is easy to see that this is (only) solved by $x(t) = c$ where c is some real constant (and t is any real number). Indeed the Fundamental Theorem of Calculus (and Mean Value Theorem) tell us that this is the general solution, thus we get a solution space of 1-dimension. This means that a single parameter dictates every possible solution, in this case that parameter is c.

The second example we look at is a slightly generalised version of this

$$x' = f(t) \tag{3.2}$$

Once again using the Fundamental Theorem of Calculus and the Mean Value Theorem, we find that the general solution is

$$x(t) = x_0 + \int_0^t f(s)dx$$

where once again our space of solution is one-dimensional, governed in this case by the constant x_0.

The third example is one where things get interesting (and also an incredibly important example).

$$x' = ax \tag{3.3}$$

In fact this is a whole family of differential equations for every $a \in \mathbb{R}$. Some minor knowledge of calculus tells us that

$$x(t) = ce^{at}$$

is a solution to this differential equation (once again c can be any real constant). However it is not immediately obvious that this is the general solution to the differential equation. Let us show that this is is the case.

Let $\tilde{x}(t)$ be any solution to the equation. Our claim is that (3.3) is a constant multiple of e^{at}. We show this by proving that the ratio of the two functions is always constant

$$\begin{aligned} \frac{d}{dt}(\tilde{x}(t)e^{-at}) &= \tilde{x}(t)(-ae^{-at}) + \tilde{x}'(t)(e^{-at}) \\ &= -a\tilde{x}(t)e^{-at} + a\tilde{x}(t)e^{-at} \\ &= 0 \end{aligned}$$

This example indeed illustrates a general principle for solving ODEs: guess and check. This is to say that it is often easier to guess an answer to an ODE and then verify that their solution works.

4 Useful Pictures

We've said many times that solving ODEs is a difficult, often impossible, task. We reiterate that here: solving ODEs is a difficult, often impossible, task. However, it is the case that the ODE can give us a lot of information about its solution(s) which we often summarise in various pictorial formats.

A differential equation gives us a way of computing the slope of the tangent to the function at any given point (this is, after all, what the derivative measures).

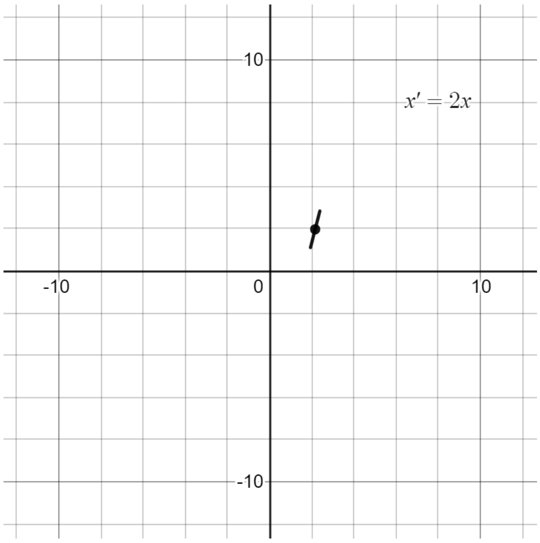

Figure 1: Drawing the slope at one point

Thus one thing we can do to try and visualise the behaviour of the function is to find its derivative at every point [2]. This called the slope field or direction field [3]

Seeing the slope field, we can turn the question of solving a differential equation to a visual one: we need to find a function which is tangent to every slope line. Of course there are a variety of possible solutions depending on where one starts from (one would hope that after deciding where to start from, there is only *one* solution. we will see that in sufficiently nice conditions, this is the case). These solutions are also called integral curves. We illustrate a few examples below in Figure 3.

We see that the slopes are all the same along any horizontal line. This is because the differential equation is independent of t. Such an ODE is called autonomous. For more general ODEs, we can still ask the question of what is set of points (t, x) such that $f(t, x) = k$ for some constant k. In this case we call the set of points *isoclines* (in particular then horizontal lines are isoclines for autonomous ODEs).

Consider the equation

$$x' = 2x$$

It is clear that the case with $x = 0$ (horizontal line in Figure 3) is a special case as this is the only constant solution. In this case we call 0 an equilibrium equilibrium or steady state or stationary point. Consider what happens we are slightly above $x = 0$. In this case as t goes to infinity (one often thinks of this time evolving), $x(t)$ gets further and further away from 0 (if this is modelling the position of a particle, then this shows that the particle grows further and further from its initial position (and at an increasing pace)). In particular if $x(t)$ it becomes more and more positive and if $x(t)$ is negative, it becomes more and more negative. We often represent this in what is called a phase line or phase portrait.

The fact that solutions near 0 move away from 0 as t increases, means that although 0 is an equilibrium point, it is an unstable one (also called a source, if one imagines this as the flow a fluid). The opposite is

[2] In practice, one often limits themselves to a finite set of points

[3] Note the multiple names which will be a common theme here. It seems that the only thing harder than solving a differential equation is naming things related to them consistently.

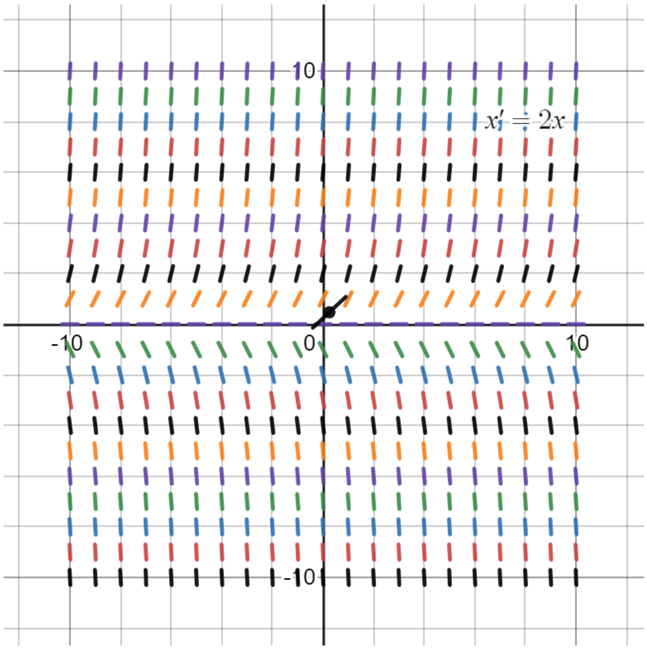

Figure 2: Drawing a slope field on a grid for $x' = 2x$

of course a stable equilibrium (or sink) where solutions within some neighbourhood of the equilibrium point tend towards the equilibrium. Such a case occurs if we consider Equation 3.3 for $a < 0$. In this case 0 is again an equilibrium point, but now it is a stable one.

Notice the difference in behaviour between $a > 0$ and $a < 0$. Behaviour for all $a > 0$ is qualitatively the same and the same holds true for $a < 0$. In such cases we say that the equation $x' = ax$ for $a > 0$ (or $a < 0$) is stable (this is different from 0 being a stable point; in this case we are talking about the family of equations being a stable one.).

It seems suggestive that we left out the case for $a = 0$ above. Indeed that is because the behaviour of the function is completely different for $a = 0$, since $x(t)$ is a constant in this case (this is exactly very example of an ODE we considered). It is at this point that 0 goes from being a source to a sink where the slightest change in a causes it to go one way or the other. We say then that $a = 0$ is a *bifurcation* in the 1-parameter family of equations $x' = ax$.

5 The Logistic Equation

Imagine you are trying to model the growth of a population. We know that if a population is small and is in ideal conditions (easily accessible food, few predators, lots of space, etc.). A population will grow exponentially. However, we also know that this cannot continue forever. As a population grows larger and larger, it will start pushing towards the limits of the available resources. In fact if the population grows too large for its environment (for example if there's not enough food or too many predators), then one would expect the population to decrease. A simple equation, known as the logistic equation[4], that models this behaviour is

$$x' = ax\left(1 - \frac{x}{N}\right) \tag{5.1}$$

where $a \in \mathbb{R}$ is the growth rate of the population and N is the carrying capcity or the ideal population size. Notice that if x is small then $x' \approx ax$ and if $x > N$ (i.e. the population is greater than the carrying capacity) then $x' < 0$. We immediately see that this equation is still autonomous (we still have no t in the equation) but it is no longer linear. This might already suggest that this is a somewhat more difficult problem than

[4] More precisely, the logistic equation is the solution to the differential equation with $N = 1$, but that's really more of a "tomato, tomato" situation

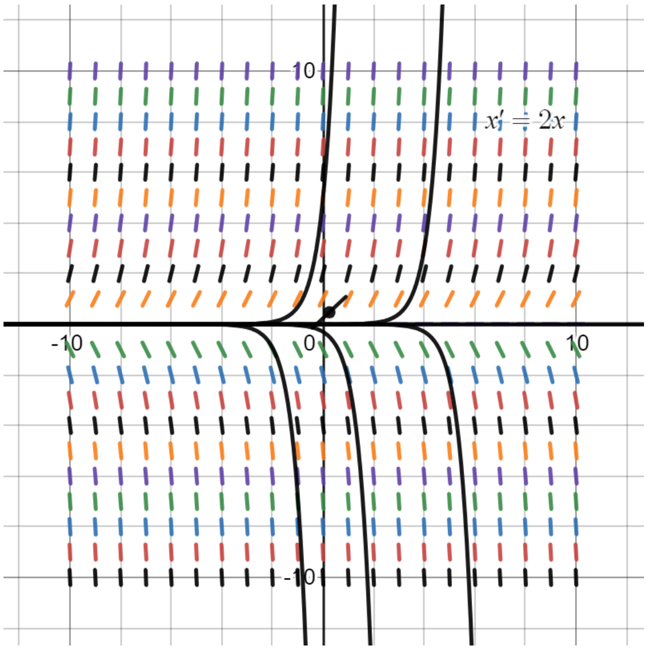

Figure 3: Slope field with some solutions (solutions illustrated in black)

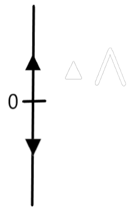

Figure 4: Phase line/portrait for $x' = 2x$

before (it is). Nevertheless, we can make our life a bit easier with one small assumption: without loss of generality we can take N to be 1 (we just choose appropriate units of x to make this work. One can think of x modelling the proportion of the ideal population rather than the actual size of the population). We thus define

$$f_a(x) = ax(1 - x) \tag{5.2}$$

We see that $f_a(x)$ is 0 for $x = 0$ and $x = 1$ (these are hence our equilibrium points), positive if $x \in (0, 1)$ and negative otherwise. This already tells us that $x \equiv 0$ is an unstable solution while $x \equiv 1$ is a stable one. We can solve for other solutions to this differential equation (for other initial conditions) using a technique known as separation of variables.

We solve the ODE using separation of variables (see Subsection 1.1).

$$
\begin{aligned}
x' &= ax(1-x) \\
\frac{x'}{ax(1-x)} &= 1 \\
\frac{1}{a}\int \frac{1}{x(1-x)}dx &= \int 1 dt \\
\frac{1}{a}\int \frac{1}{x} + \frac{1}{1-x}dx &= t + c \\
\frac{1}{a}\ln\left|\frac{x}{1-x}\right| &= t + c \\
\left|\frac{x}{1-x}\right| &= e^{at+ac} \\
\frac{x}{1-x} &= c_2 e^{at} \\
x &= \frac{c_2 e^{at}}{1 + c_2 e^{at}}
\end{aligned}
$$

Thus our solution is

$$x(t) = \frac{c_2 e^{at}}{1 + c_2 e^{at}} \tag{5.3}$$

5.1 Parameterising the general solution

Recall that that we also had special solutions to the equation, namely $x \equiv 1$ and $x \equiv 0$. We then see that $c_2 = 0$ recovers one of the solutions! Unfortunately the same is not true for the other solution (we would require $c_2 = \infty$). We can divide the numerator and denominator by c_2 to recover $x \equiv 1$ but we now lose $x \equiv 0$. This suggests that there might be a better parameterisation of the solutions. We see that

$$x_0 := x(0) = \frac{c_2}{1 + c_2}$$

We can solve for c_2 and substitute that into Equation 5.3 to get

$$x(t) = \frac{x_0 e^{at}}{1 - x_0 + x_0 e^{at}} = \frac{x_0}{(1 - x_0)e^{-at} + x_0} \tag{5.4}$$

With this parametrisation, $x_0 = 0$ and $x_0 = 1$ get us the two special solutions.

Let's analyse this equation to see what we can learn of it. We will consider the case $a > 0$ since it is clear from the differential equation that cases for $a < 0$ are simply going to be reflections of the positive case (besides it is remarkably rare for populations to have a negative growth rate). Suppose the initial point x_0 is between 0 and 1. In this case, x' is positive so x will increase as t increases (as we already predicted) and x' will tend toward 0. However, looking at Equation 5.4 we can see that no value of t will make $x(t) = 1$. Hence $x = 1$ forms an asymptote. Suppose x_0 is greater than 1. Then x' is negative. In this case, as before, x will continually decrease, approaching $x = 1$ but never intersecting it. The most interesting case is when $x_0 < 0$. In this case, we know that $x(0) = x_0 < 0$. Note that at $t = 0$ the denominator starts of as a positive number, 1. However as t grows larger and larger, $(1 - x_0)e^{at}$ will tend towards 0. Since x_0 is a fixed negative number, this means the denominator will be 0 at some point in time. In fact, we can work quite easily that this occurs at $t = \frac{1}{a}\ln(\frac{-x_0}{1-x_0})$. This means that in this case the function will blow up to $-\infty$ in finite time.[5]

6 Linear System of ODEs

A differential equation of the form

$$X' = A(t)X + f(t)$$

[5]We let the reader decide how they feel about negative population sizes exploding to $-\infty$.

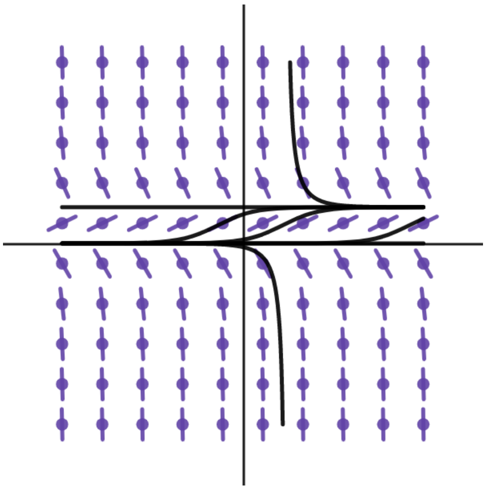

Figure 5: Solutions to the logistic equation (solutions in black)

where $A(t)$ is an $n \times n$ matrix and f is a map from an interval $I \subset \mathbb{R}$ to $\mathbb{R}^n$ is known as a linear system of ODEs. $A(t)$ is called the matrix of coefficients (its entires being the coefficients) and the function f is called the inhomogeneity. If $f = 0$ then the equation is called homogeneous and if $A(t)$ is a constant matrix then we have what is called a constant coefficient ODE. As usual, if $X' = F(X)$ for some F, then all X_0 that satisfy $F(X_0) = 0$ are called equilibrium points of a system (in the case of a linear system, $F(X) = A(t)X + f(t)$, but this statement applies more generally).

We will start by considering the simplest case of a linear system: a homogeneous, constant coefficient ODE (as we will see solving a homogeneous ODE allows us to solve the general system, see Subsection 6.2). In this case $X_0 = 0$ is always a equilibrium point. If $\det(A) = 0$ then we have a space of equilibrium. Reducing further to case of A being 2×2, if $\det(A) = 0$ we have a straight line of equilibrium points in $\mathbb{R}^2$ (we of course ignore the uninteresting case of $A = 0$).

As mentioned, we will start by simply considering the case in $\mathbb{R}^2$. In other words our differential equation is of the form

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} \tag{6.1}$$

The system would be easy to solve if $b = 0$ and $c = 0$ as we would left with equations of the form $x' = ax$ and $y' = dy$ which already know how to solve. This occurs if A is a diagonal matrix. If A is diagonalizable, then we can change our coordinates to make A diagonal and solve the system. In either case, we more or less get $x = x_0 e^{at}$ and $y = y_0 e^{dt}$ as our general solutions. Moreover, in the one-dimensional, case the system would would look like $x' = ax$, which again has the solution $x(t) = x_0 e^{at}$. This might inspire us to define the answer for the 2×2 case (and the general case) to be

$$X(t) = e^{At}$$

Despite the nonsense that this looks, there is a way of interpreting what it means to "exponentiate a matrix", using the Taylor expansion of e^x. In other words, we define

$$e^{At} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = I + tA + \frac{t^2}{2} A^2 + \frac{t^3}{3!} A^3 + \dots$$

Of course this is an infinite series so we need to decided whether it converges or not (and what convergence even means in this case), but at least each of the terms in the series makes sense.

As we will see later, this is in fact always convergent and does indeed solve our ODE and forms a general solution to our ODE. Moreover, we have that

$$\frac{d}{dt}e^{At} = Ae^{At}$$

Another educated guess one may make, once again looking at the one-dimensional case, is that the answer will be of the form

$$x(t) = e^{\lambda t} v$$

where λ is a real number and v is a (constant) vector in $\mathbb{R}^n$, which are parameters to be determined. Assuming this to be case, we can substitute this in Equation 6.1 to conclude that

$$\lambda e^{\lambda t} v = A(e^{\lambda t} v)$$

for all t. This implies that $Av = \lambda v$ or in other words that λ is an eigenvalue with eigenvector v.

6.1 Example

Let us try our hand with an example (with the second guess for now). Suppose we are given

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \underbrace{\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix} \tag{6.2}$$

Using our favourite method of finding eigenvalue/eigenvector pairs, we determine that the eigenvalues of A are 3 and -1 with eigenvectors $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ respectively. We then have two solutions

$$X_1(t) = e^{3t}\begin{pmatrix} 3 \\ t \end{pmatrix}$$
$$X_2(t) = e^{-t}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

However observe that any linear combination of X_1 and X_2 is also a solution. This leads us to superposition principle (also known as the linearity principle).

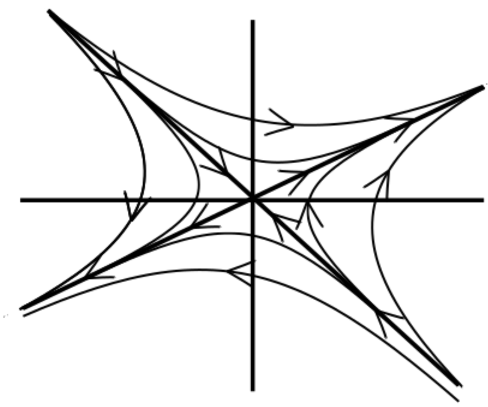

Figure 6: Phase Portrait of Equation 6.2

6.2 Superposition/Linearity Principle

Suppose $X_1(t)$ is such that it solves

$$X' = A(t)X + f_1(t)$$

and $X_2(t)$ solves

$$X' = A(t)X + f_2(t)$$

Then for real a_1, a_2, $X(t) = a_1X_1(t) + a_2X_2(t)$ solves

$$X' = A(t)X + a_1f_1(t) + a_2f_2(t)$$

This is easily verified by substituting the solution into the differential equation (and using the fact that multiplication with a matrix is linear). A consequence of this is that solutions to a homogeneous, linear system of ODEs forms a vector space.

Another consequence is the fact that the general solution $X'(t) = A(t)X + f(t)$ is given by

$$X(t) = X_{genhom}(t) + y(t)$$

where $X_{genhom}(t)$ is the general solution to the homogeneous system of equations $X' = AX$ and $y(t)$ is one *particular* solution to $X'(t) = A(t)X + f(t)$ (recall the similarities to solving a linear system of equations given by $AX = b$, where the solution is given by $b + \tilde{X}$ where $\tilde{X}$ is the space of solutions solves $AX = 0$ (if b lies in the range of A)).

This means that if X maps to $\mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$ then the space of solutions to $X' = AX$ forms a $n-$dimensional vector space. In fact if $X_1(t), \dots, X_n(t)$ are (linearly independent) solutions to $X' = AX$ then

$$X(t) = a_1X_1(t) + \dots a_nX_n(t) \quad a_1, \dots, a_n \in \mathbb{R}$$

is the general solution to $X' = AX$. This in fact holds in general but we will only prove this for the case when A can be diagonalised (for now). Later we will show the general case.

Lemma 6.1 *The general solution to $X' = AX$ where X maps to $\mathbb{R}^n$ and A is a diagonalisable $n \times n$ matrix is given by*

$$X(t) = a_1X_1(t) + \dots a_nX_n(t)$$

where the X_i themselves are linearly independent solutions (meaning $X_1(t), \dots, X_n(t)$ are linearly independent for all t) to the system of equations.

Proof. Let $v_1, \dots, v_n$ be eigenvectors of A that form a basis for $\mathbb{R}^n$. Suppose their respective eigenvalues are $\lambda_1, \dots \lambda_n$.

Suppose the initial value problem is given by

$$\begin{cases} X' = AX \\ X(0) = a_1v_1 + \dots a_nv_n \end{cases} \tag{6.3}$$

Then

$$Y(t) = a_1e^{\lambda_1 t}v_1 \dots a_ne^{\lambda_n t}v_n$$

solves this initial value problem.

Suppose $Z(t)$ is another solution to this IVP. Since $v_1, \dots v_n$ is a basis for $\mathbb{R}^n$, we can write $Z(t) = b_1(t)v_1 + \dots b_n(t)v_n$ where the b_i are real-valued functions. We know that $b_i(0) = a_i$ since $Z(0) = Y(0) = X(0)$. Also note that

$$\begin{aligned} Z'(t) &= b_1'(t)v_1 + \dots b_n'(t)v_n \\ AZ(t) &= A(b_1(t)v_1) + \dots b_n(t)v_n) \\ &= b_1\lambda_1v_1 + \dots + b_n\lambda_nv_n \end{aligned}$$

Equating coefficients of the v_i (which are uniquely determined since the v_i for a basis) we get that $b_i'(t) = \lambda_ib_i(t)$ and $b_i(0) = a_i$. We know that for each i, this is uniquely solved by $b_i(t) = e^{\lambda_i t}a_i$ implying that $Z = Y$. This allows us to conclude that the solutions to $X' = AX$ are uniquely determined by the initial value. □

Remark 6.2. We know that that b_i above are differentiable as they are the composition of two differentiable functions: Z and the linear projection onto v_i which is a linear map with constant coefficients/entries (hence in particular is differentiable with respect to t).

6.3 Types of Systems

We can categorise the different systems of equation based on the eigenvalues.

6.3.1 Saddle Point

Suppose we are given

$$X' = \underbrace{\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}}_{A} X \tag{6.4}$$

Then we know that its eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -1$ with eigenvectors $v_1 = (3, 1)$ and $v_2 = (1, -1)$ respectively. Note that this is a case where the eigenvalues are of opposite sign. In this case the phase portrait would look like Figure 7.

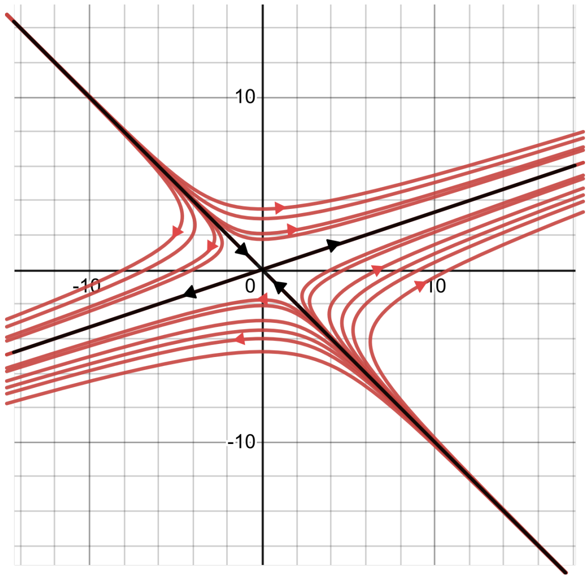

Figure 7: Saddle point

Note that as time as evolves points on the line spanned by v_2 move towards the origin (exponentially) while points on the line spanned by v_1 move away (also exponentially). In this case we call v_2 the stable line and v_1 the unstable line. Now consider a point that is not on either of these lines but is some linear combination of v_1 and v_2. In this case as time evolves, the component for v_2 will shrink to 0 while the component for v_1 will shoot off to infinity leading to this above shape. This is called having a saddle point at 0. Any time we have eigenvalues of opposite sign we get something like above.

6.3.2 Unstable Node

The next question, of course, is of course is what happens if we have two eigenvalues of the same sign. We first consider the case of both eigenvalues being positive. As an example, we can consider

$$X' = BX \tag{6.5}$$

where $B = A + 2I$. Then the eigenvalues of B are $\lambda_1 = 5$ and $\lambda_2 = 1$ with v_1 and v_2 as before. In this case everything moves away from the origin giving us what is called an "unstable node at 0". However note that since the λ_1 is greater, the v_1 component of any point will increase much faster as $t \to \infty$. So eventually the paths will seem parallel to v_1. Conversely as $t \to 0$, the v_1 component will also decrease to 0 faster than v_2 so the integral curves (aka solutions) become tangent to v_2 as t approaches 0. Hence we conclude that the phase portrait will look like so.

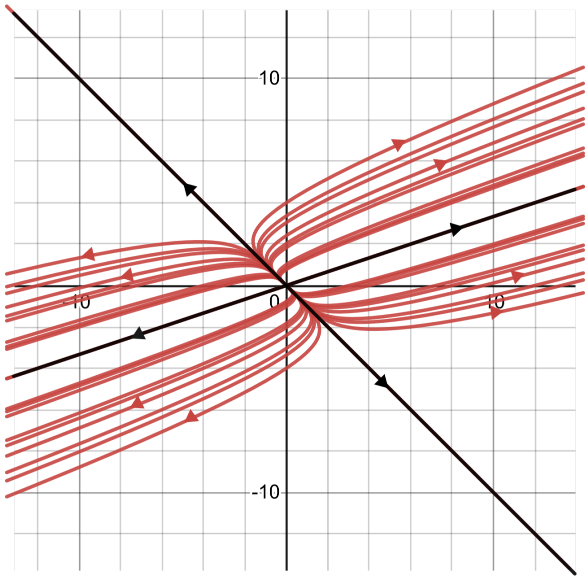

Figure 8: Unstable Node

6.3.3 Stable Node

We next look at the case when both eigenvalues are negative. Perhaps unsurprisingly, the picture will be quite similar to the previous with some slight modifications. Continuing our tradition of having a concrete example, we consider

$$X' = CX \tag{6.6}$$

where $C = A - 5I$. Our eigenvectors remain v_1 and v_2 as usual but their corresponding eigenvalues are now $\lambda_1 = -2$ and $\lambda_2 = -6$. Now as we evolve time, everything will approach the origin. Thus we call this situation having a "stable node at 0". However now v_2 approaches the origin faster than v_1 so the the integral curves will reflect this by becoming tangent to v_1 as $t \to \infty$.

6.3.4 Center

Of course not every real matrix will have eigenvalue. More particularly, it may not have *real* eigenvalues but it will certainly have *complex* eigenvalues. Indeed eigenvalues (for real matrices) come in conjugate pairs. However we can close our eyes, pretend everything is real, and in the end split things into their real and complex complex components to get genuinely real solutions. Let us demonstrate what this means. Suppose we are given

$$X' = \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{J} X \tag{6.7}$$

In this case, the characteristic equation is $p(\lambda) = \lambda^2 + 1$. The roots of this polynomial are i and $-i$ which are thus our complex eigenvalues. As mentioned, we don't worry about these being complex and proceed as normal. We see that

$$J - iI = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}$$

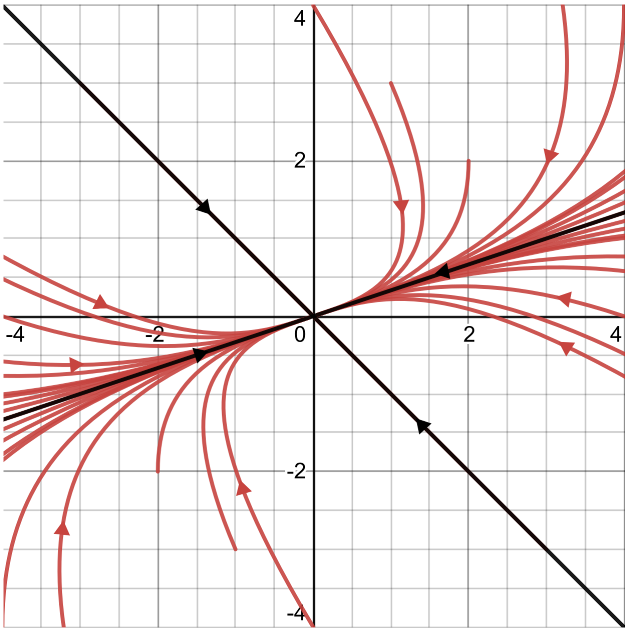

Figure 9: Stable Node

and that the vector $v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$ lies in its kernel hence is the corresponding eigenvector for $\lambda_1 = i$. Similarly we conclude $v_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$ is the eigenvector with eigenvalue $\lambda_2 = -i$ (in fact this can be concluded without any calculations. If A is a matrix with real entries and has an eigenvector v with eigenvalue λ, then $\overline{v}$ is an eigenvector with eigenvalue $\overline{\lambda}$ where $\overline{v}$ is defined in the obvious way: taking the complex conjugate of each entry). Thus our two solutions are

$$z_1(t) = e^{it}v_1, z_2(t) = e^{-it}v_2$$

Expanding this using Euler's formula, we get

$$\begin{aligned} e^{it}v_1 &= (\cos t + i\sin t)\begin{pmatrix} i \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} i\cos t - \sin t \\ \cos t + i\sin t \end{pmatrix} \\ &= \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} + i\begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \end{aligned}$$

We claim that

$$X_1(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}, X_2(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

are (linearly independent) solutions to Equation 6.7 (note we would have gotten the same solutions if we had chose to expand z_2 instead).

It is easy to see that the solutions will travel in a circular path: counterclockwise as time moves forward and clockwise as time moves backward. This is known has having a center at 0. This occurs anytime the eigenvalues are purely imaginary numbers.

A few claims were made up in the above example. Let us prove them formally.

Lemma 6.3 *Suppose $A \in \mathbb{R}^{n\times n}$. Suppose v is an eigenvector with a eigenvalue λ. Then*

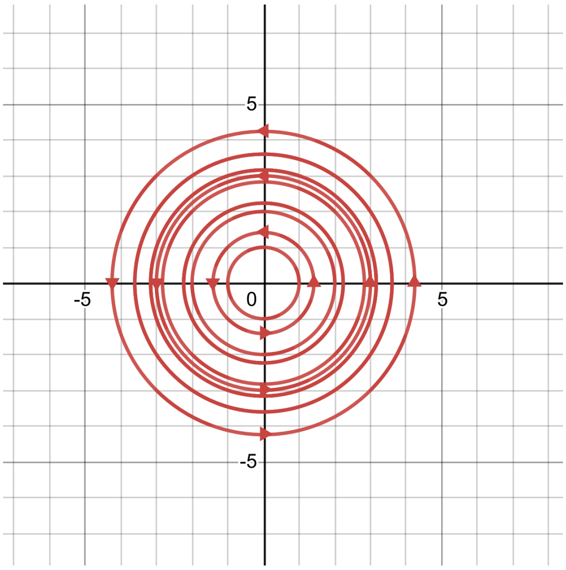

Figure 10: Center at 0

- *$\overline{\lambda}$ is an eigenvalue of A with eigenvector $\overline{v}$*
- *If λ is not real, then v is not in $\mathbb{R}^n$ (it has complex entries). Moreover, $Re(v)$ and $Im(v)$ are linearly independent.*

Proof. The first statement is easily verified by noting that $\overline{A} = A$ since A has real entries. Therefore

$$Av = \lambda v \Leftrightarrow \overline{Av} = \overline{\lambda v} \Leftrightarrow A\overline{v} = \overline{\lambda}\overline{v}$$

In order to verify the second statement suppose $v = u + iw$ where $u, w \in \mathbb{R}^n$. We will prove that u and w are linearly independent via contradiction. So suppose there exists real number s and t and some $v_0 \in \mathbb{R}^n$ such that $u = sv_0$ and $w = tv_0$. Then $v = u + iw = (s + it)v_0$. Since v_0 is a multiple of v, it must also be an eigenvector of A with eigenvalue λ. This means that

$$Av_0 = \lambda v_0$$

The left side of this equation is always in $\mathbb{R}^n$ however if λ is not real then λv_0 will not be. Thus we get a contradiction if λ is non-real. This shows that u and w are linearly independent so neither can be 0, thus v must have complex entries. □

Lemma 6.4 *$Z(t)$ is a complex solution to $X' = AX$ (where A is a real matrix) if and only if $Re(Z(t))$ and $Im(Z(t))$ are also solutions.*

Proof.

$$Z'_{Re}(t) + iZ'_{Im}(t) = Z'(t) = AZ(t) = AZ_{Re}(t) + iZ_{Im}(t)$$

□

Next we want to show that we have found the general solution to Equation 6.7.

Lemma 6.5 *The general solution Equation 6.7 is given by $x(t) = a(-\sin t, cost) + b(\cos t, \sin t)$*

Proof. Suppose $y(t) = (u(t), v(t))$ is another solution to the differential equation. Let $f(t) = (u(t) + iv(t))e^{-it}$. Then

$$\begin{aligned} f'(t) &= (u'(t) + iv'(t))e^{-it} - ie^{-t}(u(t) + iv(t)) \\ &= (-v(t) + iu(t))e^{-it} + e^{-it}(-iu(t) + v(t)) \\ &= 0 \end{aligned}$$

Therefore $y = \alpha e^{it}$ where α is some complex number implying that y is a linear combination of $X_1(t)$ and $X_2(t)$ as given above. □

6.3.5 Spiral

Suppose we have

$$X' = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} X \tag{6.8}$$

We find that the eigenvalues are $\lambda_1 = 2 + i$ and $\lambda_2 = 2 - i$ with corresponding eigenvectors $v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$. Thus we find that a solution is given by

$$\begin{aligned} Z(t) &= e^{(2+i)t}\left(i\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) \\ &= e^{2t}(\cos t + i\sin t)\left(i\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) \\ &= e^{2t}\begin{pmatrix}-\sin t\\ \cos t\end{pmatrix} + ie^{2t}\begin{pmatrix}\cos t\\ \sin t\end{pmatrix} \end{aligned}$$

Thus the two solutions are given by the real and the imaginary parts:

$$X_1(t) = e^{2t}\begin{pmatrix}-\sin t\\ \cos t\end{pmatrix}, X_2 = e^{2t}\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}$$

The general solution is then of course some linear combination of X_1 and X_2.

Consider $X_1(t)$. As t increases the $(-\sin t, \cos t)$ component makes the point go around in the origin (with period 2π) while the e^{2t} causes the magnitude to increase. Thus we get a spiraling out. In order to determine the direction of the spiral (i.e. clockwise or counterclockwise) we can either investigate $(-\sin t, \cos t)$ (as t increases the $x-$coordinate decreases while the $y-$coordinate increases, therefore counterclockwise) or we can try a point and determine which way the the tangent vector points. For example substituting $X = (1,0)$ into Equation 6.8 we find that $X' = (2,1)$ implying that that the spiral must be going counterclockwise.

6.4 Repeated eigenvalues (Part I)

Suppose we have

$$X' = \begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix} X \tag{6.9}$$

where we will denote matrix by A, as usual. The characteristic polynomial of A is $\lambda^2 - 2\lambda + 1$ which has a repeated root for $\lambda = 1$. Hence 1 is an eigenvalue with algebraic multiplicity 2. However $A - I$ is a matrix of rank 1, hence there is only one eigenvector namely $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. This is gives one solution with

$$X_1(t) = ce^t\begin{pmatrix}1\\2\end{pmatrix}$$

However we expect a two dimensional solution space (a statement which we will justify later) hence we know there should be one other linearly independent solution.

From the study of linear algebra, we recall in this case A must have a basis in generalised eigenvectors. In this we can compute the generalised eigenvector of A is $w = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Hence we might expect the second

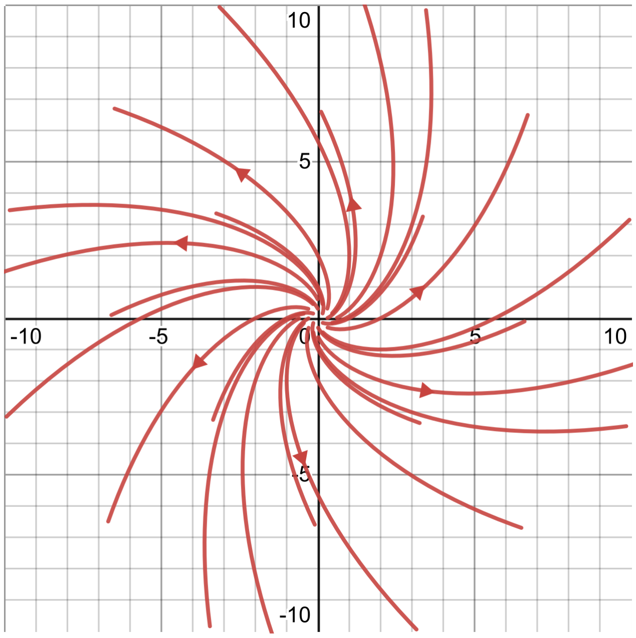

Figure 11: Spiral (counterclockwise)

solution to be of the form $X(t) = \alpha(t)v + \beta(t)w$ with $\beta \neq 0$ (the case with $\beta = 0$ is covered by X_1 above). Assuming that $X(t)$ does solve our differential equation we see that

$$\begin{aligned} X'(t) &= A(\alpha(t)v + \beta(t)w) \\ &= \alpha(t)Av + \beta(t)Aw \\ &= \alpha(t)\lambda v + \beta(t)(\lambda w + v) \\ &= (\lambda\alpha(t) + \beta(t))v + \lambda\beta(t)w \end{aligned}$$

On other hand we also know that

$$X'(t) = \alpha'(t)v + \beta'(t)w$$

equating coefficients (since (v, w) is a basis, the coefficients are unique), we get the following system of equations

$$\begin{cases} \alpha' = \lambda\alpha + \beta \\ \beta' = \lambda\beta \end{cases}$$

The second equation is easily solved by $\beta(t) = ce^{\lambda t}$ where c can be any constant (in fact, as we know, the equation is *only* solved by this). Recall that we are trying find a *particular* solution to the differential equation (one that is linearly independent of X_1). Hence, for simplicity, we can take $c = 1$ above giving us $\beta(t) = e^{\lambda t}$.

This, however, leaves us to solve to solve for α. One might expect that we can simply use the theory built up so far since this is simply a homogeneous, linear system of equations. However, one can check that this system is exactly the case we are considering right now: the case with repeated eigenvalues with a basis in generalised eigenvectors. Therefore we will have to resort to some combination of guesswork and being clever.

In this case, we recall that the homogeneous equation $\alpha' = \lambda\alpha$ would be solved by $\alpha(t) = ce^{\lambda t}$ where c is a constant. One might guess that if instead we vary the constant in some appropriate manner (i.e. make it a function of t), we might be able to solve for $\alpha'(t) = \lambda\alpha(t) + e^{\lambda t}$. Therefore suppose $\alpha(t) = y(t)e^{\lambda t}$. Assuming this to be a solution, we get

$$\alpha'(t) = y'(t)e^{\lambda t} + \lambda y(t)e^{\lambda t}$$

It is then clear that if $y'(t) = 1$ then we have a solution. In other words we can take $y(t) = ct$ for any constant c. Once again we are only interested in one particular solution. Therefore, we take the simplest one to conclude that $\alpha(t) = te^{\lambda t}$. We then finally have a second, linearly independent solution to our ODE:

$$X_2(t) = te^{\lambda t} v + e^{\lambda t} w$$

In other words the general solution to the ODE is given by

$$X(t) = a_1 X_1(t) + a_2 X_2(t)$$

where a_1, a_2 are constants determined by the initial conditions, as usual.

Let us substitute the specific values of λ, v and w to get a clearer picture of the solutions (everything thus far will hold for any repeated eigenvalues with a basis in generalised eigenvectors). We have that our solution is given by

$$\begin{aligned} X(t) &= a_1 e^t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_2 \left(te^t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + e^t \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right) \\ &= a_1 e^t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_2 e^t \left(t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right) \\ &= (a_1 e^t + a_2 te^t) \begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_2 e^t \begin{pmatrix} 0 \\ -1 \end{pmatrix} \end{aligned}$$

We see that as $t \to \infty$ the component corresponding to $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ increases much more quickly that its partner, implying that the solution curves become increasing parallel to the line spanned by $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. On the other hand as $t \to -\infty$, this component also approaches 0 more quickly, so in the limit the solution curves becomes tangent to $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ at the origin. If the eigenvalue were negative, then things remain essentially identical except the arrows are reversed.

6.5 Repeated Eigenvalues (Part II)

It's possible that one has repeated eigenvalues *and* a basis of eigenvectors. An example is

$$X' = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} X \tag{6.10}$$

where the standard basis vectors are eigenvectors themselves. In this case every vector is an eigenvector. Although the example might seem like a special case, it in fact highlights the general case since the assumptions on the matrix mean that it is a multiple of the identity. If it's a positive multiple of the identity (like the example above), then everything tends away from the origin in the direction parallel to itself. Note how extremely unstable this is since the mildest of perturbations cases the long term behaviour to be completely different. If we have a negative multiple of the identity, then the situation reverses and everything approaches the origin (and of course this is a very stable situation: regardless of where you start, you tend towards the origin).

7 Trace-Determinant Plane

We are interested in classifying the different dynamical systems, roughly based on what the phase portraits look like. For example, all centers are roughly the same (the only thing that varies is the frequency and the direction of the rotation); all saddle points are essentially the same (up to some rotation and stretching), etc. What we realise is that almost all of this information is determined by the eigenvalues and in particular by the sign of the eigenvalues, whether they are real or complex, etc. If we can then work out the relationships between the eigenvalues (ideally without solving for them), we can classify the dynamical systems relatively easily.

This is where we remember from our study of linear algebra that the eigenvalues of a matrix can by found by looking at the roots of the characteristic polynomial of the matrix and in the 2×2 case, this polynomial is completely determined by the trace and determinant. Indeed we have that the characteristic polynomial p_A of a 2×2 matrix A is given by

$$p_A(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A)$$

We know then, for example, that if the discriminant of the above quadratic is positive, then we have 2 distinct real eigenvalues. If in addition the determinant is positive, then we know the eigenvalues share sign. Finally we can use the trace to determine what exactly their sign is. This gives us a near complete description of the qualitative behaviour of A. Repeating this analysis for the other cases, we summarise our findings below.

Eigenvalues	Normal form	How to detect	Shape
$\lambda_1 > \lambda_2 > 0$	$\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$	$\mathrm{Tr}(A)^2 > 4\det(A)$	Unstable node
$\lambda_1 > 0 > \lambda_2$	$\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$	$\det(A) < 0$	Saddle
$0 > \lambda_1 > \lambda_2$	$\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$	$\det(A) > 0$, $\mathrm{Tr}(A) < 0$	Stable node
$\lambda = \alpha + i\beta$, $\alpha > 0$	$\begin{pmatrix}\alpha & -\beta\\ \beta & \alpha\end{pmatrix}$	$\mathrm{Tr}(A) > 0, 4\det(A) > \mathrm{Tr}(A)^2$	Spiral source
$\lambda = \alpha + i\beta, \alpha < 0$	$\begin{pmatrix}\alpha & -\beta\\ \beta & \alpha\end{pmatrix}$	$\det(A) > 0, \mathrm{Tr}(A) < 0$ $4\det(A) < \mathrm{Tr}(A)^2$	Spiral sink

Table 1: Generic cases for $X' = AX$

These are the generic cases (to defined more precisely later), but roughly speaking, this is what 'most' 2×2 matrices look like: they either have 2 distinct eigenvalues or 2 complex eigenvalues (which are conjugates). This can be verified by the fact that the above cases have covered almost all of the trace-determinant plane.

There do remain, however, some degenerate or exceptional cases that we still need to verify. In general these occur when we have an equality of some kind (in other words something is equal to 0 and as one can imagine this is rarely a good thing). These are summarised in Table 2.

We say a bifurcation occurs in an ODE when change the parameters a bit causes the qualitative behaviour to change drastically (for example changing from a source to a sink). This is a definition we will make more precise later. These information is often summarised in a bifurcation diagram. Note that the trace-determinant plane is an example of a bifurcation diagram where the axes and curves are used to group similarly behaved equations together and bifurcations occur when we go from one region to another.

8 Canonical Forms and Genericity

Definition 8.1 (Hyperbolicity)**.** The origin is called a hyperbolic equilibrium for the ODE $X' = AX$ if all eigenvalues of A have non-zero real part.

In the case when A is a 2×2 matrix, hyperbolicity holds whenever $\det(A) < 0$ or if $\det(A) > 0$ and $\mathrm{Tr}(A) \neq 0$ (see trace-determinant plane).

Definition 8.2 (Genericity)**.** A property (for example of matrices) is called generic if it is satisfied on a dense, open subset (of $\mathbb{R}^{n\times n}$ for example).

Eigenvalues	Normal form	How to detect	Shape
$\lambda = i\beta$	$\begin{pmatrix} 0 & -\beta \\ \beta & 0 \end{pmatrix}$	$\mathrm{Tr}(A) = 0, \det(A) > 0$	Center
$\lambda_1 = \lambda_2 > 0$	$\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$	$\mathrm{Tr}(A)^2 = 4\det(A)$, $\mathrm{rank}(A - \lambda_1 I) = 1, \mathrm{Tr}(A) > 0$	Unstable, source
$\lambda_1 = \lambda_2 < 0$	$\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$	$\mathrm{Tr}(A)^2 = 4\det(A)$, $\mathrm{rank}(A - \lambda_1 I) = 1, \mathrm{Tr}(A) < 0$	Stable, sink
$\lambda_1 = \lambda_2 > 0$	$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}$	$A = \lambda_1 I, \mathrm{Tr}(A) > 0$	(Very) unstable, source
$\lambda_1 = \lambda_2 < 0$	$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}$	$A = \lambda_1 I, \mathrm{Tr}(A) < 0$	Stable, sink
$\lambda_1 > 0, \lambda_2 = 0$	$\begin{pmatrix} \lambda_1 & 0 \\ 0 & 0 \end{pmatrix}$	$\det(A) = 0, \mathrm{Tr}(A) > 0$	Unstable
$\lambda_1 < 0, \lambda_2 = 0$	$\begin{pmatrix} \lambda_1 & 0 \\ 0 & 0 \end{pmatrix}$	$\det(A) = 0, \mathrm{Tr}(A) < 0$	Stable
$\lambda_1 = \lambda_2 = 0$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\mathrm{Tr}(A) = \det(A) = 0$, $\mathrm{rank}(A) = 1$	Unstable
$\lambda_1 = \lambda_2 = 0$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$A = 0$	Constant

Table 2: Degenerate cases for $X' = AX$

Theorem 8.3 *The property of having n distinct eigenvalues is generic for n × n matrices. In other words the subset of n × n matrices which have n distinct eigenvalues is open and dense in* $\mathbb{R}^{n\times n}$.

Corollary 8.3.1 (Cayley Hamilton Theorem) *Let A be an* $n \times n$ *matrix and let* $p_A(\lambda)$ *be its characteristic polynomial. Then* $p_A(A) = 0$.

Proof. The statement is easily verified for diagonal matrices (the diagonal entries are the eigenvalues and also the zeroes to the characteristic polynomial). Also note that $p_S(T^{-1}ST) = T^{-1}p_S(S)T$ for every invertible $T \in \mathbb{R}^{n\times n}$ and every $S \in \mathbb{R}^{n\times n}$. Thus the statement holds true not only for diagonal matrices but for diagonalisable matrices as well. Since the subset of matrices with n distinct eigenvalues is dense in $\mathbb{R}^{n\times n}$, there exists a sequence of diagonalisable matrices $(A_i)_{i=1}^{\infty}$ (with n distinct eigenvalues) that converges to A. In particular this means that the entries of the sequence of the matrices approach the entries of A. The coefficients of the characterisitic polynomial of a matrix depend continuously on the entries of the matrix. Thus as $A_i \to A$ we have that $p_{A_i} \to p_A$. Since $p_{A_i}(A_i) = 0$ we get $p_{A_i}(A_i) \to p_A(A) = 0$. □

8.1 Canonical Forms

Suppose we are studying our good old friend $X' = AX$. Suppose we set $X = TY$ where T is some invertible matrix. Then

$$Y' = (T^{-1}X)' = T^{-1}X' = T^{-1}AX = T^{-1}ATY$$

Defining $C := T^{-1}AT$ we get another differential equation $Y' = CY$. Importantly if we choose T cleverly we can make C a 'simple' matrix so that $Y' = CY$ is easily solved. Once we have a solution for Y, we can easily find X, the solution to our original differential equation, since by definition $X = TY$. What does it mean for C to be simple? Ideally we would want it to be diagonal of course. But this of course not always possible. Thus we go for the next best thing: the Jordan Canonical Form (JCF). Technically even this is not always possible over the reals. However, what we can put the matrix into JCF as if we were over the complex numbers and use the complex solutions to get (pairs of) real solutions, as we've done before.

The Jordan Canonical form allows to perform a change of basis to write every matrix as a block matrix of the form

$$C = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}, J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

where each J_i is an $l_i \times l_i$ matrix. The J_i's are called Jordan blocks. Note that a diagonal matrix is a special case of the above form where each Jordan block is 1×1.

Remark 8.4. The same eigenvalue could appear in different Jordan block. The total number of times that λ appears in C (along the diagonal of course) is exactly the algebraic multiplicity of λ.

Remark 8.5. If $\lambda \in \mathbb{C}$ is an eigenvalue, we know that $\overline{\lambda}$ is as well. This also means that $\overline{J}$ is a Jordan block in A.

8.1.1 Example

Suppose we know that a 5×5 matrix has an eigenvalue λ that has algebraic multiplicity 5 as well. Here are some possible Jordan blocks.

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & & & & \\ & \lambda & & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

The question then is how do we know which of the Jordan blocks we have? This takes a bit more work.

Note that any Jordan block of size l can be written as

$$J = \lambda I + N$$

where N is a nilpotent matrix (this means that $N^l = 0$ but $N^{l-1} \neq 0$). As a consequece then we can distinguish which of the canonical forms by looking at the kernel of $(A - \lambda)$ raised to various power. To be specific, $\dim\ker((A - \lambda I)^m)$ is the number of Jordan blocks of size less than or equal to m. In first case given above, we see that $\dim\ker(A - \lambda) = 1$. This is enough to characterise this matrix completely since this is the only way to a 5×5 Jordan block with one eigenvector. However for the middle two examples, we see that $\dim\ker(A - \lambda I) = 2$ for both of them (they both have two eigenvectors). We then try the next power: $\dim\ker(A - \lambda I)^2$ is 3 for the first matrix (second in row) and 4 for the other matrix (third in row).

8.2 Solving Jordan Blocks

The question now boils down to how to solve a system of the form $Y' = CY$ where C is in canonical form and possibly complex. One thing to note is that Jordan decomposition allows us to break the space into invariant subspaces (that is after all one of the motivations for the decomposition). What this means is that solutions for the Jordan blocks can be found independently and 'stacked' up together (think of the analogy with diagonal matrices where we get n distict equations that can be solved independenly and when combined give us a solution to the entire system). Thus we we only need solutions to $Y' = JY$ where the matrix J is of the form

$$J=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1\\ & & & \lambda\end{pmatrix}=\lambda I+\underbrace{\begin{pmatrix}0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1\\ & & & 0\end{pmatrix}}_{N}$$

Suppose $Y(t)=e^{t\lambda}Z(t)$ is a solution to $Y'=JY$. Then

$$\begin{aligned}Z'(t)&=-\lambda e^{-t\lambda}Y+e^{-t\lambda}Y'\\&=-\lambda e^{-t\lambda}Y+e^{-t\lambda}(\lambda Y+NY)\\&=N(e^{-t\lambda}Y)\\&=NZ\end{aligned}$$

Thus we solve for $Z'=NZ$, where N is a nilpotent matrix as shown above. This is equivalent to writing

$$\begin{pmatrix}z_1\\ \vdots\\ z_{n-1}\\ z_n\end{pmatrix}'=\begin{pmatrix}z_2\\ \vdots\\ z_n\\ 0\end{pmatrix}$$

Since $z_n'=0$, we know that $z_n(t)=c_n$ where c_n is some arbitrary constant. Then $z_{n-1}=c_nt+c_{n-1}$ where c_{n-1} is again some real constant. Continue this way we get $z_1(t)=p(t)$ where p is some polynomial of degree $n-1$ (since n-th derivative is 0). This is the general solution (which makes sense, the space of polynomials of degree at most $n-1$ is of dimension n). Then $Y(t)=e^{t\lambda}Z(t)$ solves $Y'=JY$, where we may split Y into its real and imaginary components if necessary. We have thus found a solution for all linear systems! A pat on the back is well-deserved but we postpone that for after a discussion of matrix exponentials.

9 Matrix Exponentials

There is a second method of reaching the same answer with a bit less work (or rather with most of the work swept under the rugs of past theorems and lemmas). We claim that a (in fact the) solution to $Y'=(\lambda I+N)Y$ (where N is nilpotent) with $Y(0)=y_0$ is given by

$$e^{\lambda I+N}y_0=e^{\lambda}t\sum_{k=0}^{\infty}\frac{t^k}{k!}N^ky_0$$

Normally we would have to worry about convergence of infinite series, however since N is nilpotent the above series is finite. Once again we have a product of $e^{\lambda}t$ with some polynomial in t as we did before (the coefficents come from y_0). This might lead one to conjecture that the general solution to $X'=AX$ for *any* A is given by

$$e^{tA}$$

where once again we are more or less using the exponential notation as shorthand for its Taylor expansion. We first need to ensure that this actually makes sense, that is we really do have convergence. We first need a norm. It is a fact that in finite dimensions that all norms (and also all inner products) are equivalent (any norm can be bounded by a multiple of another norm and similarly with inner products). Thus we can quite frankly choose any norm we want. We will choose one that is quite common in this setting, called the operation norm which is defined as

$$\|A\|:=\sup_{\|v\|\le 1}\|Av\|=\sup_{v\in\mathbb{R}^n,v\neq 0}\frac{|Av|}{|v|}$$

We may equivalently define it as the largest singular value of A (in the appropriate setting).

By definition we have the fact that $\|A\| \le \|A\| \|v\|$. This leads to the lovely fact that $\|AB\| \le \|A\| \|B\|$ since

$$\|AB\| = \sup_{\|v\|\le 1} \|ABv\| \le \|A\| \sup_{\|v\|\le 1} \|Bv\| \le \|A\| \|B\|$$

Remark 9.1. $\mathbb{R}^{n\times n}$ with this norm is called a Banach algebra.

Proposition 9.2 *The series*

$$e^{tA} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

converges absolutely.

Proof. We need to show that

$$\sum_{k=0}^{\infty} \left\| \frac{t^k}{k!} A^k \right\|$$

is finite. With our previous statement this is easy to see, since

$$\begin{aligned} \sum_{k=0}^{\infty} \left\| \frac{t^k}{k!} A^k \right\| &= \sum_{k=0}^{\infty} \frac{|t|^k}{k!} \left\| A^k \right\| \\ &= \sum_{k=0}^{\infty} \frac{|t|^k}{k!} \|A\|^k \\ &= e^{|t|\|A\|} \end{aligned}$$

□

We also have the following lovely statements about matrix exponentials. Note that the first property is often called the semigroup property and the latter property is an if and only if.

Lemma 9.3 *Let $s, t \in \mathbb{R}$ and $A, B \in \mathbb{R}^{n\times n}$. Then*

1. $e^{(t+s)A} = e^{tA} e^{sA}$
2. $e^{t(A+B)} = e^{tA} e^{tB}$ *if* $AB = BA$

Proof. For the first statement (and for the second statement, the proofs are near identical), we see that

$$\begin{aligned} e^{(t+s)A} &= \sum_{k=0}^{\infty} \left(k! \sum_{i+j=k} \frac{s^i}{i!} \frac{t^j}{j!} \right) \frac{A^k}{k!} \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} \frac{s^i}{i!} A^i \frac{t^j}{j!} A^j \right) \\ &= \left(\sum_{i=0}^{\infty} \frac{s^i}{i!} A^i \right) \left(\sum_{j=0}^{\infty} \frac{t^j}{j!} A^j \right) \\ &= e^{sA} e^{tA} \end{aligned}$$

The same proof works for the second case since we can use the binomial theorem again as the matrices A, B commute. □

A consequence of the second property is that the exponential of a matrix is always invertible (we take $B = -A$).

Finally we want to make a comment on the differentiability of this map to make sure that it is what we expect.

Lemma 9.4

$$\frac{d}{dt}e^{tA} = Ae^{tA}$$

Proof.

$$\begin{aligned}
\frac{d}{dt}e^{tA} &= \lim_{h\to 0}\frac{\exp((t+h)A)-\exp(tA)}{h} \\
&= \lim_{h\to 0}\frac{\exp(tA)\exp(hA)-\exp(tA)}{h} \\
&= \exp(tA)\lim_{h\to 0}\frac{\exp(hA)-I}{h} \\
&= \exp(tA)\lim_{h\to 0}\frac{1}{h}\left(hA+\frac{h^2}{2}A^2+\frac{h^3}{3!}A^3+\dots\right) \\
&= \exp(tA)A
\end{aligned}$$

□

Much like the exponential function with real numbers, matrix exponentiation gives us a unique solution to a differential equation. In particular, $X(t) = e^{tA}$ is the unique solution to the initial value problem $X' = AX$ with $X(0) = X_0$. The proof is the exact same as with real numbers: suppose $Z(t)$ is another solution. Define $W(t) = Z(t)e^{-tA}x_0$. Then

$$\begin{aligned}
W'(t) &= \frac{d}{dt}(Z(t)e^{-tA}x_0) \\
&= (-Z(t)Ae^{-tA})x_0 + (Z'(t)e^{-tA})x_0 \\
&= (-Ae^{-tA}Z(t) + e^{-tA}AZ(t))x_0 \\
&= 0
\end{aligned}$$

(we use the fact that A commutes with its exponential). Therefore $W(t)$ is a constant x_0. Since $Z(0) = e^{0A}x_0 = x_0$, Z and X agree everywhere and we are done.

10 Inhomogeneous Linear Systems

Suppose we have an equation of the form

$$X' = Ax + f(t) \tag{10.1}$$

Then we know from the superposition principle (see Subsection 6.2) that the general solution to this system is given by

$$X(t) = y(t) + e^{tA}v$$

where $y(t)$ is a particular solution to the ODE and v is some arbitrary vector that is determined by the initial conditions. Thus our goal is to find just *one* solution to this ODE. This leads us to Duhamel's Principle.

10.1 Duhamel's Principle

We will guess that $y(t) = e^{tA}v(t)$ is a solution where $v(t)$ is a function to be determined (this is the technique of variation of constants). Assuming y is a solution we can plug this in Equation 10.1 to find the left and right hand sides are

$$y' = Ae^{tA}v(t) + e^{tA}v'(t) \tag{LHS}$$

$$Ay + f(t) = Ae^{tA}v(t) + f(t) \tag{RHS}$$

Equating the two we find that

$$v'(t) = e^{-tA}f(t)$$

Hence by the Fundamental Theorem of Calculus we find that

$$v(t) = \int_0^t e^{-sA} f(s) ds$$

(recall we only need a particular solution so we can ignore the constant of integration by setting it to 0). Therefore our particular solution y is given by

$$y(t) = e^{tA} v(t) = e^{tA} \int_0^t e^{-sA} f(s) ds = \int_0^t e^{(t-s)A} f(s) ds$$

Remark 10.1. Note that $t - s$ is always positive. Although not particularly relevant to this example, this is an important note in other contexts.

Remark 10.2. What we mean by integrating a vector-valued function is to integrate each of the component functions and 'stack' them together to get another vector.

11 Linearisation

The idea with linearisation is to simply use the framework we've Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth vector field and we have the differential equation

$$x' = f(x) \tag{11.1}$$

Although it's hard to give an explicit solution to this, we can still try and determine it's qualitative behaviour, at least locally. Let $p \in \mathbb{R}^n$ be arbitrary. If $f(p)$ is non-zero, then for q near p we are going to have that $f(q)$ is close to $f(p)$. Therefore the flow is going to look like (almost) parallel lines in this neighbourhood. If $f(p)$ is 0, we need to do a bit more work. One thing we can do is consider the Taylor expansion of f. We know that

$$f(x) = f(p) + f'(p)(x-p) + \underbrace{O(|x-p|)}_{\text{error}}$$

Since we assume $f(p) = 0$, the first term disappears and by substituting $y = x - p$, Equation 11.1 becomes

$$y' = Df(p)y + O(|y|)$$

Since $Df(p)$ is a linear map, this looks just like the linear equations we have studied thus far, except there is the added error term. The hope is that this error terms is going to be small so by studying the linear system

$$y' = Df(p)y$$

we can get a pretty good idea of how the true system behaves. For example if there are sources or sinks in the linearised system, they will also appear in the true system. The phase portraits will also be similar (similarity will be defined more precisely later) if the origin is hyperbolic (recall this means that $\Re(\lambda) \neq 0$ for every eigenvalue λ of $Df(p)$).

This determines a procedure that we can use to study such equations

1. Find the steady states/equlibria (the points where $x' = f(x)$ is 0). These are points where the solutions are constant and don't change over time.
2. Linearise near these steady states
3. Tie everything up into a big picture

The cryptic 'tie everything up into a big picture' can be best illustrated with an example.

11.1 Example

We will consider the case of swinging a pendulum, kind of. More precisely we will look at the case when a mass is attached to a (rigid) rod allowed to swing freely in a vertical circle (we use a rod instead of a string because we don't want to worry about cases where the string may fold onto itself or something).

The equation modelling this situation is given by

$$mx'' + rx' = -c\sin x$$

where x is the angle made by the rod with the vertical. Here m is the mass of the mass (names are difficult for physicists), r is the constant of friction (therefore $r \geq 0$) and c is some arbitrary constant (the exact details, such as length of the rod, strength of gravity, etc, are used to set c). To make the analysis a bit simpler we will assume m and c to be 1. Then in particular we have the equation

$$x'' + rx' + \sin x = 0 \tag{11.2}$$

We can use our standard trick to convert this second order equation to a system of first order equation: let $v = x'$. Then we have

$$\begin{cases} x' &= v \\ v' &= -rv - \sin x \end{cases}$$

We first find the the steady states or in other words where v is 0 (this implies that v' is 0 since $v' = \frac{df(x(t))}{dt} = f'(x)x'(t) = f'(x)v = 0$. These are the points where $\sin x = 0$ or in other words where $x = k\pi, k \in \mathbb{Z}$. k being even corresponds with the mass hanging on the bottom and k being odd is when the mass is at the top in a perfectly vertical position. Simple intuition tells us that the the former equlibria should be stable (at least if $r > 0$)and the latter should be unstable. Let us see if the equations agree with this. First we see that

$$\begin{pmatrix} x \\ v \end{pmatrix}' = f(x,v), f(x,v) = \begin{pmatrix} v \\ -rv - \sin x \end{pmatrix}$$

Then

$$\begin{aligned} Df(k\pi,0) &= \begin{pmatrix} 0 & 1 \\ -\cos x & -r \end{pmatrix}\Bigg|_{x=k\pi} \\ &= \begin{cases} \begin{pmatrix} 0 & 1 \\ 1 & -r \end{pmatrix}, & k \text{ odd} \\ \begin{pmatrix} 0 & 1 \\ -1 & -r \end{pmatrix}, & k \text{ even} \end{cases} \end{aligned}$$

Let us consider the case with k odd first. In this case the determinant of the matrix is -1 so the eigenvalues are real and of opposite sign. We know this will result in a saddle and hence be unstable. In fact around these points we expect the phase portrait to (roughly) have a saddle as well. This lines up with our intuition above.

Now let us consider the case with k even. Then we know the eigenvalues are

$$\lambda_{1,2} = \frac{-r \pm \sqrt{r^2 - 4}}{2}$$

If $0 < r < 2$ then we will have complex eigenvalues implying that we will have a spiral. Since the real part is positive, solutions are going to spiral in and by looking at the first column we can even infer that spiral is going to be clockwise. This corresponds with the angle tending towards 0 and its speed decreasing, as we would expect the pendulum to behave. If $r > 2$ then we get 2 real eigenvalues, both of which are negative. In this case all equilibria will be stable. This corresponds with the friction becoming so strong that the pendulum can actually become stuck and 'stable' at odd angles.

We get some rather interesting behaviour at $r = 0$, in the frictionless case. In this case we get a center which again should make sense. If there is no friction, then the pendulum continues swinging on its path ad infinitium.

12 Exact Differential Equations

Suppose we have an ODE of the form

$$P(x,y)dx + Q(x,y)dy = 0$$

is exact if there exists a function $f(x,y)$ such that

$$\frac{\partial f}{\partial x} = P(x,y), \frac{\partial f}{\partial y} = Q(x,y)$$

The general solution to such ODEs is given by the one-parameter family

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

It would be nice to know when an ODE is exact. Luckily we have the following theorem which not only tells us when an equation is exact but even gives a construction for the function f.

Theorem 12.1 *Suppose we have an ODE of the form*

$$P(x,y)dx + Q(x,y)dy = 0$$

where P, Q are defined on a simply connected region and are C^1. Then the equation is exact if and only if

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

Proof. The forward direction follows from a theorem in analysis. Namely if a function f is smooth (or even just C^1) then

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Thus we need show the reverse direction. Thus we assume that

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

Suppose such an f were to exist. Then we know that f must satisfy

$$\frac{\partial f}{\partial x} = P(x,y) \Rightarrow f(x,y) = \int_{x_0}^{x} P(x,y)dx + R(y)$$

Typically we get a constant term when integrating, but in this case the constant term depends on y, hence why R becomes a function of y. By assumption the derivative of f with respect to y is Q. Thus we get that

$$\begin{aligned}
Q(x,y) = \frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\left(\int_{x_0}^{x} P(s,y)ds + R(y)\right) \\
&= \int_{x_0}^{x} \frac{\partial}{\partial y} P(s,y)ds + R'(y) \\
&= \int_{x_0}^{x} \frac{\partial}{\partial x} Q(s,y)ds + R'(y) \\
&= Q(x,y) - Q(x_0,y) + R'(y)
\end{aligned}$$

where of course x_0 is some arbitrary constant in the domain. We then conclude that $R'(y) = Q(x_0,y)$. Therefore

$$R(y) = \int_{y_0}^{y} Q(x_0,s)ds$$

(we ignore the integration constant for now since account for it later when giving the general solution to the ODE). Therefore

$$f(x, y) = \int_{x_0}^{x} P(s, y)ds + \int_{y_0}^{y} Q(x_0, s)ds$$

All that remains to check is that this f does indeed satisfy the conditions. Now it is clear that

$$\frac{\partial f}{\partial x} = P(x, y)$$

The other one takes slightly more work

$$\begin{aligned}
\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\int_{x_0}^{x} P(s, y)ds + Q(x_0, y) \\
&= \int_{x_0}^{x} \frac{\partial}{\partial y} P(s, y)ds + Q(x_0, y) \\
&= \int_{x_0}^{x} \frac{\partial}{\partial x} Q(s, y)ds + Q(x_0, y) \\
&= Q(x, y) - Q(x_0, y) + Q(x_0, y) \\
&= Q(x, y)
\end{aligned}$$

□

12.1 Example

Suppose we have the equation

$$(2x + y\cos x)dx + (2y + \sin x - \sin y)dy = 0 \tag{12.1}$$

In this case we have $P(x, y) = 2x + y\cos x$ and $Q(x, y) = 2y + \sin x - \sin y$. Since

$$\frac{\partial P}{\partial y} = \cos x = \frac{\partial Q}{\partial x}$$

we can conclude using the previous theorem that the differential equation is exact. Thus we can construct f as given by the proof as well. We start with

$$f(x, y) = \int P(x, y)dx + R(y) = x^2 + y\sin x + R(y)$$

Then we know that

$$\begin{aligned}
\frac{\partial f}{\partial y} &= Q(x, y) \\
\sin x + R'(y) &= 2y + \sin x - \sin y
\end{aligned}$$

Therefore

$$R(y) = y^2 + \cos y$$

Finally we conclude that the general solution to Equation 12.1 is

$$f(x, y) = x^2 + y\sin x + y^2 + \cos y = c$$

12.2 Integrating Factors

Sometimes an equation is not exact, but we can make it exact by multiplying it with a function. Such a function is called the integration factor. As an example consider the equation

$$(t^2x - t)dx + xdt = 0 \tag{12.2}$$

Suppose there was some function $h(t)$ (we assume h is a function of t because we know this works. In general it's difficult to know what the integration factor should be a function of). Then we would have

$$(h(t)(t^2x-t))dx+(xh(t))dt=0$$

which we assume to be exact. This means that

$$\begin{aligned}\frac{\partial}{\partial t}(h(t)(t^2x-t)) &= \frac{\partial}{\partial x}(xh(t))\\ h'(t)(t^2x-t)+h(t)(2tx-1) &= h(t)\\ \frac{h'(t)}{h(t)} &= \frac{2-2tx}{t^2x-t}\\ &= -\frac{2}{t}\end{aligned}$$

Since the right hand side is a function of t, we can integrate to conclude that

$$\log(h(t))=-2\log t$$

or in other words that

$$h(t)=\frac{1}{t^2}$$

Similar manipulations can be performed if h is function of $x, xt, \frac{x}{t}, \frac{t}{x}$, etc. Often it is not obvious what h should be a function and requires some trial and error. Hence why although we know that any ODE $x'=f(x,t)$ with $f\in C^1$ in a neighbourhood of (x_0,t_0) admits an integration factor, in principle this is a useless technique because it is incredibly hard to find the integration factor.

However there do exist some special cases where integration factors *can* be found and these serve as useful examples. Suppose we have an equation of the form

$$\frac{dy}{dx}+P(x)y=Q(x) \tag{12.3}$$

Since the derivative and y are raised to the power of 1, we call this a linear differential equation and specifically a linear differential equation of order 1. One way of solving this equation is to solve the homogeneous version (i.e. set $Q=0$) using separation of variables and then solve the ODE using variation of constants. A second method, however, is to use integration factors.

We can first rewrite the equation to get

$$[P(x)y-Q(x)]dx+dy=0$$

Suppose we had an an integrating factor $u(x)$ (in this case we know that the integrating factor is always a function of the dependent variable, in this case x). Then we would have

$$\begin{aligned}\frac{\partial}{\partial y}[u(x)P(x)y-u(x)Q(x)] &= \frac{\partial}{\partial x}u(x)\\ u(x)P(x) &= u'(x)\\ \frac{u'(x)}{u(x)} &= P(x)\\ u(x) = e^{\int P(x)dx}\end{aligned}$$

Note that when integrating the exponent we don't need to worry about the integration constant since we don't need the general solution. Substituting this integration factor into Equation 12.3 we get

$$\begin{aligned}e^{\int P(x)dx}\frac{dy}{dx}+P(x)e^{\int P(x)dx}y &= e^{\int P(x)dx}Q(x)\\ \frac{d}{dx}(ye^{\int P(x)dx}) &= e^{\int P(x)dx}Q(x)\\ y &= e^{-\int P(x)dx}\left(\int e^{\int P(x)dx}Q(x)\right)+ce^{-\int P(x)dx}\end{aligned}$$

We apologise for the horror we have bestowed upon the reader.

12.3 Bernoulli Equations

Bernoulli equations are slight variations of the linear differential equations seen before which aren't themselves linear equations but can be turned into ones. Specifically, a Bernoulli equation is of the form

$$\frac{dy}{dx} + P(x)y = Q(x)y^n \tag{12.4}$$

In the case of $n = 1$ we can solve by separation of variables so let us assume that $n \neq 1$. Then we multiply both sides with $(1-n)y^{-n}$ to get

$$(1-n)y^{-n}\frac{dy}{dx} + (1-n)y^{1-n}P(x) = (1-n)Q(x)$$

Now if we define $u = y^{1-n}$ we see the above equation becomes

$$\frac{du}{dx} + (1-n)P(x)u = (1-n)Q(x)$$

which is a linear differential equation of order 1 we know how to solve.

13 Dynamical Systems

Consider our goto differential equation:

$$X' = AX$$

where A is some fixed matrix. We know the general solution is given by

$$X(t) = e^{tA}v$$

where v is determined by the initial conditions. There are two ways we may wish to study this equation. We could fix v and consider how the equation behaves as we vary t. This gives us exactly the trajectory or orbit of a solution.

On the other hand, we could fix t and consider what happens as we vary v. Such a map is often denoted Φ_t is known as the *flow* of the system (this function maps some initial point to where it would be sent by the solution at time t). What we get then is a map from v to $e^{tA}v$, which in this case is simply a linear isomorphism on R^n. We in fact have a family of isomorphisms $\{e^{tA}\}_{t\in\mathbb{R}}$ that satisfy the following property: $e^{0A} = \mathrm{Id}$ and $e^{(t+s)A} = e^{tA}e^{sA}$. A family of bijective maps $\{\phi_t\}_{t\in\mathbb{R}}$ on $\mathbb{R}^n$ that satisfies these properties (i.e. $\phi_0 = \mathrm{Id}$ and $\phi_{s+t} = \phi_s \circ \phi_t$) is called a dynamical system. In different areas of math, we make different assumptions on the properties that the ϕ_t must satisfy. For our purposes, we assume them all to be smooth. In principle we index simply over the positive reals or the natural numbers (which would be like looking at discrete time). If we were feeling particularly adventurous, we could replace $\mathbb{R}^n$ with a manifold instead. But for now, we leave things as they are.

It is perhaps useful to consider why the 'semigroup property' of dynamical systems (i.e. that $\phi_{s+t} = \phi(s) \circ \phi(t)$ is important. The claim is that with this property, by only looking at the initial conditions, we can in fact study all solutions. This is because the semigroup property allows us to 'stitch together' solutions in a certain sense. Suppose we have a solution that begins at x and passes through y_1 in time t. Suppose we have another solution that starts at y_1 and passes through y_2 in time s. Then one would hope that the original solution that began at x also passes through y_2, but at time $t+s$ (and in fact hopefully the path from x to y_2 would be given by combining/stiching together the two solutions in the obvious manner, should this be possible). This is exactly the property characterised by the semigroup property.

14 Existence and Uniqueness of solutions

The existence and uniqueness of solutions is of course a key point with differential equations. We have already said that most differential equations can't be solved explicitly. Part of the problem is that there may be no analytic solution. However, the situation may be even more dire than that: there may genuinely be no solution. Suppose we are given that

$$x' = \begin{cases} -1, & x \geq 0 \\ 1, & \text{otherwise} \end{cases}$$

Consider the situation at $x = 0$. At this point, the point 'wants' to decrease due to it negative gradient. However as soon as it does so, the derivative becomes positive causing it to move up. Hence why the above differential equation has no solutions. We can argue more precisely using a statement in analysis which tells us that the derivative of a function can never have a jump discontinuity (roughly speaking a jump discontinuity means that the left and right hand limits of a function differ, however the derivative of a function at a point is a limit so we know if it exists, the left and right hand limits are the same). Hence we know that if a solution to $x' = f(x)$ is to exist then f must at the very least be continuous.

Sometimes solutions may exist but need not be unique. Suppose we have

$$x' = 3x^{2/3}$$

with initial condition $x_0 = 0$. Then $x(t) = 0$ and $x(t) = t^3$ are both solutions that satisfy this initial value problem. The problem here is that f is continuous but not smooth (or even differentiable everywhere). In fact what we will see is that given $x' = f(x)$, we are guaranteed to have unique solutions to the ODE, if f is continuously differentiable. However, it is hard to ensure that we have a solution that is defined everywhere as is illustrated by equation

$$x' = x^2$$

whose general solution is given by

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

We see that this is not defined at $t = x_0^{-1}$. This leads us to the local uniqueness and existence theorem.

Theorem 14.1 (Local Existence and Uniqueness) *Suppose we are considering the ordinary differential equation $X' = F(X)$ where $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable. Then for all $a \in \mathbb{R}^n$ there exists an interval $I = (\alpha, \beta)$ containing the origin such that $X' = F(X)$ has a unique solution $X : I \to \mathbb{R}^n$ satisfying $X(0) = a$.*

Although we will get on to proving this theorem soon enough, let us consider some of its consequences. For example, if we have two solutions where the corresponding intervals intersect, is it the case that the solutions agree on the intersection? The answer is yes. This follows from uniqueness of the solutions: take a point in the intersection and use this to define our initial condition. We know that both solutions solve the differential equation, therefore must be equal (on the intersection). This suggests that we can 'patch' together local solutions (as discussed previously) to get solutions defined on a larger interval. By pasting together all the local solutions, we can get a maximal interval of existence (to be precise, the interval can be found by taking union of the intervals guaranteed by the above theorem as we range over all the point in $\mathbb{R}^n$).

14.1 Preliminary Theory

There are several results we need before we can prove the local existence and uniqueness theorem. On account of the author being too small brain, we largely only include the definitions and statements of the theorems we need and omit the proofs.

Because we will refer to it many times, we also include the definition of uniform continuity and uniform convergence.

Definition 14.2 (Uniform Continuity)**.** A function $f : E \to \mathbb{R}^m$ is said to be uniformly continuous if for every $\epsilon > 0$ we can find a $\delta > 0$ such that given $x, y \in E$ satisfying $|x - y| < \delta$, we have $|f(x) - f(y)| < \epsilon$ (in particular the choice of δ is independent of x and y).

Definition 14.3. A sequence of functions $\{f_n : E \to \mathbb{R}^m\}$ is said to converge uniformly to a function $f : E \to \mathbb{R}^m$ if for every $\epsilon > 0$ there exists some $N \in \mathbb{N}$ such that

$$\sup_{x \in E} |f_n(x) - f(x)| < \epsilon$$

for every $n \geq N$.

Recall that if a sequence of continuous functions converges uniformly then the function they converge to is also continuous.

Definition 14.4 (Uniformly Bounded)**.** Let $\{f_\alpha\}$ be a family of functions where each f_α is a map from $E \subset \mathbb{R}^n$ to $\mathbb{R}^m$. Suppose there exists some M such that $|f_\alpha(x)| < M$ for all $x \in E$ and all α. Then we say that the $\{f_\alpha\}$ are uniformly bounded.

Definition 14.5 (Equicontinuous)**.** Let $\{f_\alpha\}$ be a family of functions where each f_α is a map from $E \subset \mathbb{R}^n$ to $\mathbb{R}^m$. The $\{f_\alpha\}$ are said to be equicontinuous if for every $\epsilon > 0$ we can find some $\delta > 0$ such that for all $x, y \in E$ and for all f_α, if we have $|x - y| < \delta$ then $|f_\alpha(x) - f_\alpha(y)| < \epsilon$ (in particular the choice of δ is independent of x, y and α).

Example 14.1. The family of functions $f_\alpha(x) = \alpha x$ for $\alpha \in \mathbb{R}$ is not equicontinuous. If we instead restrict α to be in $[3, 5]$ (or any bounded interval) then the family of functions is equicontinuous. ■

Example 14.2. The family of functions $f_n(x) = x^n$ defined on the unit interval $[0, 1]$ is not equicontinuous. ■

Example 14.3. A family of Lipschitz functions which share the same Lipschitz constant is equicontinuous. ■

Example 14.4. Any finite family of uniformly continuous functions is equicontinuous. ■

We can define the sup norm on a function space where if $f : E \to \mathbb{R}^m$ then

$$\|f\|_\infty = \sup_{x \in E} |f(x)|$$

The space of continuous functions on $[0, 1]$ is often denoted $C([0, 1])$. We claim that that $C([0, 1])$ equipped with $\|\cdot\|_\infty$ is a complete metric space. This is exactly the statement that the uniform convergence of a sequence of continuous functions is also continuous.

Theorem 14.6 *Suppose $f_n : E \to \mathbb{R}^n$ is a sequence of continuous maps. If E is compact and and the $\{f_n\}$ converge uniformly, then the $\{f_n\}$ are uniformly bounded and equicontinuous.*

Proof. Let f denote the limit of the f_n. We first show uniform boundedness. Since each f_n is continuous and E is compact, we know there exists some constant M_n such that $|f_n(x)| \leq M_n$ (indeed we can take M_n to be $\|f_n\|_\infty$). Since f is also continuous, it is also going to be bounded by some M_0.

The idea is that eventually all the f_n are going to be quite close to f therefore we should able to bound (bind?) all but finitely many of them with M_0 (or something close to it at least). Then we are only left with finitely many bounds so the maximum of all of these should bound all the f_n. Let us formalise this.

Let N be such that for all $n \geq N$ we have that $\|f - f_n\|_\infty < 1$ (this is equivalent to saying that $|f(x) - f_n(x)| < 1$ for all $x \in E$). The existence of such an N is guaranteed by uniform continuity. Let $M = \max\{M_0 + 1, M_1, \ldots, M_N\}$. We claim that M is a bound for all the f_n. Clearly this holds true for $n \leq N$. Suppose $n > N$. Then

$$\|f_n\|_\infty \leq \|f_n - f\|_\infty + \|f\|_\infty < 1 + M_0$$

Thus the f_n are uniformly bounded.

For equicontinuity, we again use the fact that we can use f to approximate all but finitely many of the f_n. In particular, we see that

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f(x)| + |f(x) - f(y)| + |f_n(y) - f(y)|$$

Suppose $\epsilon > 0$ is given. Then there exists some $N \in \mathbb{N}$ such that $\|f - f_n\|_\infty < \epsilon$ for $n \geq N$. Additionally since f is continuous on a compact space, it is in particular uniformly continuous. Thus there exists a δ_0 such that if $|x - y| < \delta_0$ then $|f(x) - f(y)| < \epsilon$. Additionally, each of the f_n are uniformly continuous as well, thus there exist similar δ_n for each of them as well. It is then easy to see that the δ we need is $\delta = \min\{\delta_0, \delta_1, \ldots, \delta_n\}$. □

We then get to the Ascoli-Arzelà theorem.

Theorem 14.7 (Ascoli-Arzelà) *Let $\mathscr{F} = \{f_\alpha : E \to \mathbb{R}^m\}$ where E is a compact subset of $\mathbb{R}^n$ be an infinite family of continuous functions that is uniformly bounded and equicontinuous. Then there exists a sequence of functions in $\mathscr{F}$ that converges uniformly on E.*

Proof. You can see my notes here or open any functional analysis textbook (I think). □

Recall that we say a set A is relatively compact if its closure is compact which (for metric spaces)[6] is the same as saying every sequence in A has a convergent subsequence (although the limit may not be in A itself). We then have the following theorem

Theorem 14.8 *Let $A \subset C(K, \mathbb{R}^m)$ (this is the set of continuous functions from K to $\mathbb{R}^m$) where $K \subset \mathbb{R}^n$ is compact. Then A is relatively compact if and only if it is uniformly bounded and equicontinuous.*

Proof. Once again, my notes above or any functional analysis textbook should work. □

We also need consider how we may extend functions defined on a subset to be defined on the whole space.

Theorem 14.9 *Let $f : \overline{B(0,r)} \to \mathbb{R}^m$ (where $B(0,r) \subset \mathbb{R}^n$) be continuous. The $\overline{f} : \mathbb{R}^n \to \mathbb{R}^m$, where*

$$\overline{f}(x) = \begin{cases} f(x) & \text{if } |x| \leq r \\ f\left(r\frac{x}{|x|}\right) & \text{if } |x| > r \end{cases}$$

is continuous as well.

Proof. Trust □

Finally there are various fixed point theorems we should be aware of, none of which we prove.

Theorem 14.10 (Banach's Contraction Mapping Theorem) *Let E be a complete metric space and let $T : E \to E$ be such that there exists some $0 \leq q < 1$ where $d(T(x), T(y)) \leq q d(x, y)$. Then T has a unique fixed point. In other words there is exactly one point $z \in E$ such that $T(z) = z$.*

Remark 14.11. The proof of this theorem is constructive. In fact the construction is such that you can start with any point in E and construct a sequence (by iteratively applying T) that converges to the fixed point.

Theorem 14.12 (Brouwer's Fixed Point Theorem) *Let $B \subset \mathbb{R}^n$ denote the closed unit ball. Then if $T : B \to B$ is continuous, it has a fixed point.*

Remark 14.13. The proof of this theorem is a bit less nice unfortunately. Although we know a fixed point exists, in general, we have no way of working out what it might be.

Theorem 14.14 (Schauder-Tychonoff Theorem) *Let B be the (closed) unit ball in $C([0,1], \mathbb{R}^n)$, equipped with the usual supremum norm (in other words this is the set of continuous functions on $[0,1]$ that take values in the unit ball in $\mathbb{R}^n$). Suppose $T : B \to B$ is continuous map where $T(B)$ is relatively compact. Then T has a fixed point.*

Remark 14.15. We know that $T(B) \subset B$. Since B is bounded, $T(B)$ is always going to be bounded. Thus, by Theorem 14.8, we can equivalently assume $T(B)$ to be an equicontinuous family of functions.

[6] For a similar statement in general topological spaces, change the word sequence to net, see Wikipedia

14.2 Existence of solutions

Given an initial value problem (IVP), there are 3 things we would like to have, the existence of a solution, the uniqueness of a solution and a continuous dependence on its parameters. We begin by showing the first of these. To be precise we want the following.

Suppose $\xi_0 \in \mathbb{R}^n$ and $f : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous. We wish to find a function $x : [t_0, t_1] \to \mathbb{R}^n \in C^1([t_0, t_1]; \mathbb{R}^n)$ such that

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

and $x(t_0) = \xi_0$.

We call this the initial value problem of course. There is an equivalent way of formulating this problem, in which we need to solve an integral equation. The statement is as follows.

Let f and ξ_0 be as above. We wish to find $x : C^1([t_0, t_1]; \mathbb{R}^n)$ such that for all $t \in [t_0, t_1]$, we have

$$x(t) = \xi_0 + \int_{t_0}^{t} f(s, x(s))ds$$

Using the fundamental theorem of calculus, we can see that the two problems are equivalent. (In other words, if we find an x that solves the integral equation, then it will be a solution to the initial value problem and vice verse). In general, we will work to solve the integral equation as integrating things makes them nicer (discontinuous functions become continuous, continuous functions become differentiable, etc.).

Theorem 14.16 (Cauchy-Peano Theorem) *Suppose $f : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and bounded by M. Suppose we are given the initial value problem*

$$\begin{cases} x'(t) = f(t, x(t)) \text{ for all } t \in [t_0, t_1] \\ x(t_0) = \xi_0 \end{cases}$$

Then the IVP has at least one solution.

Proof. As mentioned previously, we will work to solve the integral equation

$$x(t) = \xi_0 + \int_{t_0}^{t} f(s, x(s))ds$$

We first define an operator $T : C^1([t_0, t_1]; \mathbb{R}^n) \to C^1([t_0, t_1]; \mathbb{R}^n)$ where

$$Ty(t) = \xi_0 + \int_{t_0}^{t} f(s, y(s))ds$$

What we wish to find, then, is some x such that $Tx = x$. As one can imagine, we will use the Schauder-Tychonoff theorem to do this.

In order to apply the theorem, we need to verify that its conditions hold. So let B be the unit ball in $C^1([t_0, t_1]; \mathbb{R}^n)$ (in other words $B = \{y \in C^1([t_0, t_1]; \mathbb{R}^n) : |y(t)| \leq 1\}$; we use C^1 because we want the solution to be C^1). We see that

$$\begin{aligned} |Tx(t)| &\leq |\xi_0| + \left| \int_{t_0}^{t} f(s, x(s))ds \right| \\ &\leq |\xi_0| + \int_{t_0}^{t} |f(s, x(s))|ds \\ &\leq |\xi_0| + M(t - t_0) \end{aligned}$$

We split the remainder of the proof into two cases. First suppose $|\xi_0| + M(t_1 - t_0) \leq 1$. Then clearly $Tx \in B$ implying that we do indeed have a map $T : B \to B$. Now suppose $t, t' \in [t_0, t_1]$. Then

$$\begin{aligned} |Tx(t) - Tx(t')| &= \left| \left(\xi_0 + \int_{t_0}^{t} f(s, x(s))ds \right) - \left(\xi_0 + \int_{t_0}^{t'} f(s, x(s))ds \right) \right| \\ &= \left| \int_{t'}^{t} f(s, x(s))ds \right| \\ &\leq M|t - t'| \end{aligned}$$

This means that $T(B)$ is a family of functions which all share the Lipschitz constant M. Therefore $T(B)$ is equicontinuous. Since we already know it to be bounded (we are assuming that $T(B) \subset B$), we have that $T(B)$ is relatively compact by Theorem 14.8. All that remains to show is that T is continuous. Since we are working in metric spaces, it suffices to show that T maps convergent sequences to convergent sequences. So suppose $\{x_k\}$ is a sequence of functions in B that converge uniformly to some function x. In other words, we have that $\|x - x_k\|_\infty \to 0$. Then

$$|Tx(t) - Tx_k(t)| \le \int_{t_0}^{t} \left|f(s,x(s)) - f(s,x_k(s))\right| ds$$

By uniform convergence, the right hand side goes to 0. This is because f is uniformly continuous (this follows from f being continuous on the compact set $[t_0,t_1] \times [-1,1]^n$. Since we are only concerned with the set B, we only need consider f restricted to this domain) thus we can get the integrand to be as small as we like given that the inputs are close enough and the inputs can be made as close as we like by uniform convergence.

To be precise, suppose $\epsilon > 0$. By uniform continuity of f, we know there is some $\delta > 0$ such that $|(s_1,x_1) - (s_2,x_2)| < \delta$ implies that $\left|f(s_1,x_1) - f(s_2,x_2)\right| < \frac{\epsilon}{t_1 - t_0}$. By uniform convergence of x_k, there is some $N \in \mathbb{N}$ such that for $k \ge N$, we have $\|x - x_k\|_\infty < \delta$. Then $|x(t) - x_k(t)| < \delta$ for all t. Thus for $k \ge N$, we have

$$\int_{t_0}^{t} \left|f(s,x(s)) - f(s,x_k(s))\right| ds < \frac{\epsilon}{t_1 - t_0}(t_1 - t_0) = \epsilon$$

With this we satisfy all the conditions of Schauder-Tychonoff which tells us there exists some $x \in B$ such that $Tx = x$, the precise statement we were aiming for.

The second case to consider is when $|\xi_0| + M(t_1 - t_0) > 1$. Let us call the quantity on the left H. Then we define

$$g(t,x) = \frac{1}{H} f(t,Hx), \eta_0 = \frac{1}{H}\xi_0$$

This gets us a new IVP which falls under case 1 and gets us a solution $y(t)$. Then $x(t) = Hy(t)$ solves the original IVP. □

There is a second proof of the Cauchy-Peano theorem that is inspired by Euler's polygonal.

Proof. Just to make notation easier, suppose we take $t_0 = 0$ and $t_1 = 1$. Let k be some natural number. Then we define

$$x_k(t) = \begin{cases} \xi_0 & \text{if } 0 \le t \le \frac{1}{k} \\ \xi_0 + \int_0^{t-\frac{1}{k}} f(s,x_k(s))ds & \text{otherwise} \end{cases}$$

Note that $x_k'(t) = f(t - \frac{1}{k}, x_k(t - \frac{1}{k}))$ which should be close $f(t, x_k(t))$ provided that k is sufficiently large. Hence the hope is that we can find a converging (sub)sequence among these x_k that will satisfy the IVP. We show that a converging subsequence exists by using Ascoli-Arzelà.

Since M is a bound for f, we see that

$$|x_k(t)| \le |\xi_0| + \int_0^{t-\frac{1}{k}} \left|f(s,x_k(s))\right| ds \le |\xi_0| + M$$

Therefore $|\xi_0| + M$ is a uniform bound for the x_k. Additionally

$$\begin{aligned} \left|x_k(t) - x_k(t')\right| &= \left|\int_{t'-\frac{1}{k}}^{t-\frac{1}{k}} f(s,x_k(s))ds\right| \\ &\le M\left|\left(t - \frac{1}{k}\right) - \left(t' - \frac{1}{k}\right)\right| \\ &= M\left|t - t'\right| \end{aligned}$$

This means that all the x_k are Lipschitz with Lipschitz constant M and hence form an equicontinuous family of functions. By Ascoli-Arzelà, we know there exists a convergence subsequence which we will denote x_k again and we denote their limit as x. All that remains to show is that x satisfies the integral equation (this ends up being easier to show than proving it satisfies the IVP). We see that

$$x_k(t) = \xi_0 + \int_0^t f(s, x_k(s))ds - \int_{t-\frac{1}{k}}^t f(s, x_k(s))ds$$

Consider what happens as we take the limit as $k \to \infty$. The left hand side goes to $x(t)$ by definition and the final term clearly goes to 0 (it bounded by $\frac{M}{k}$). Since the convergence of the x_k is uniform, taking the limit commutes with integration allowing us to conclude that

$$x(t) = \xi_0 + \int_0^t f(s, x(s))ds$$

$\square$

Consider again the initial value problem

$$x'(t) = f(t, x(t))$$
$$x(t_0) = \xi_0$$

We will try find an interval $I = [t_0, t_0 + h]$ (for $h > 0$) and a solution $x \in C^1(I; \mathbb{R}^n)$ such that x is a solution to the IVP. In this case, we will say that x is a solution to IVP$_+$. For now we only focus on moving time forward, the theory for working backwards in time in near identical and will be discussed later.

If x is a solution to IVP$_+$ on I, then we will say it can be continued to the right if there exists a pair $(\overline{x}, \overline{I})$ such that $\overline{I} \supset I$ and $\overline{x}|_I = x$. We will say a continuation is strict if $\overline{I} \supsetneq I$. This allows us to define a preorder, namely

$$(x_1, I_1) \geq (x_2, I_2) \Leftrightarrow I_1 \supset I_2 \text{ and } x_1|_{I_2} = x_2$$

By Zorn's lemma, there exists a maximal continuation (x^*, I^*) of (x, I). Note that I^* is a union of all intervals on which we have continuations of x.

Theorem 14.17 *Suppose we are given $(t_0, \xi_0) \in \mathbb{R} \times \mathbb{R}^n$. Let $A = [t_0 - h, t_0 + h] \times \overline{B(\xi_0, a)}$ for some $h, a > 0$. Let $f : A \to \mathbb{R}^n$ be continuous function bounded by M. Then the IVP*

$$x'(t) = f(t, x(t))$$
$$x(t_0) = \xi_0$$

has at least one solution x defined on

$$I = [t_0 - \min\left\{h, \frac{a}{M}\right\}, t_0 + \min\left\{h, \frac{a}{M}\right\}]$$

Moreover, any solution to the IVP defined on $J \subset I$, where J is a neighbourhood of t_0 (i.e. J contains an open set that in turn contains t_0), can be continued to a solution on I.

Remark 14.18. Although the first statement is quite similar to Cauchy-Peano, note that f is now local in space (we use a closed ball rather than all of $\mathbb{R}^n$). Hence we will require a further bit of argument.

Proof. By Theorem 14.9, we can extend f to $\overline{f}$ defined on $[t_0 - h, t_0 + h] \times \mathbb{R}^n$. By definition of the extension, $\left\|\overline{f}\right\|_\infty \leq M$ (which is to say its values are contained in the cube of 'radius' M). By Cauchy-Peano, there exists a solution $\overline{x}$ to the IVP

$$x'(t) = \overline{f}(t, x(t))$$
$$x(t_0) = \xi_0$$

Unfortunately $\overline{x}$ is not a solution to our original IVP since it may take values outside of A. However this can be easily fixed by using the fact that $\overline{x}$ is continuous. Since $\overline{x}$ is continuous, there exists $j > 0$ (where $j \leq h$) such that if $t \in [t_0, t_0 + j]$ then $|\overline{x}(t) - \overline{x}(t_0)| = |\overline{x}(t) - \xi_0| \leq a$.

Hence we have a solution to the given IVP on $I = [t_0, t_0 + j]$ for some j. Now we want to show that given a solution on any $J \subset I$, the solution can be extended to a solution on I (where of course J must be a neighbourhood of t_0).

Let (x, J) be a solution to the IVP where $J = [t_0, t_1]$ for some $t_1 \in [t_0, t_0 + h]$ (as mentioned, we will only consider the case of moving forward in time for now). We extend it as described above to get a maximal solution (i.e. defined on the maximal time interval). Then we need show that if (x^*, J^*) is this maximal solution, then $J^* \supset I$.

First we claim that J^* contains its right endpoint (in other words J^* is a closed interval). Let $t', t'' \in J^*$ with $t' < t''$. Then, by considering the integral equation instead of the IVP, we get that

$$\begin{aligned}\left|x^*(t'') - x^*(t')\right| &\le \left|\int_{t'}^{t''} f(s, x^*(s))\,ds\right| \\ &\le \int_{t'}^{t''} \left|f(s, x^*(s)\right|\,ds \\ &\le M(t'' - t')\end{aligned}$$

This means that x^* is uniformly continuous (indeed it is even Lipschitz). In particular then x^* maps a Cauchy sequence to a Cauchy sequence. By constructing a Cauchy sequence that converges to $t_1 := \sup(J^*)$, we can either define x^* on it or verify that it is continuous at that point. In principle we also need to check that $x^*(t_1) \in \overline{B(\xi_0, a)}$. However this is clear since the set is closed (and even compact), so it contains its limit points (and by the construction given, we know that $x^*(t_1)$ is indeed a limit point).

Remark 14.19. This means that any solutions (at least the ones obtained via this method) are defined on closed intervals rather than open/half-open intervals.

Then we show that $t_1 \ge t_0 + \min\left\{h, \frac{a}{M}\right\}$. If $t_1 = t_0 + h$, we are done, so suppose $t_1 < t_0 + h$. We will show that $t_1 \ge t_0 + \frac{a}{M}$ in this case. First we note that we must have $x^*(t_1)$ on the boundary of $\overline{B(\xi_0, a)}$ (if $x^*(t_1)$ was in the interior, we could run the proof again with $x^*(t_1)$ as our initial starting point and obtain a solution on $[t_1, t_1 + b]$ for some $b > 0$. We could then combine these solutions to get a solution on $[t_0, t_1 + b]$, contradicting J^* being the maximal interval). This means that

$$\begin{aligned}a &= |x^*(t_1) - \xi_0| \\ &= |x^*(t_1) - x^*(t_0)| \\ &\le (t_1 - t_0)\left\|x'\right\| \\ &\le M(t_1 - t_0)\end{aligned}$$

The third line follows from the mean value theorem in higher dimensions (see Wikipedia). Therefore we get that

$$t_1 \ge \frac{a}{M} + t_0$$

All that remains to show then is that we can continue to the left of t_0 in a similar manner. Of course we could simply run through the proof again changing the signs where necessary to obtain solutions to the left of t_0. However, if we are clever (and mathematicians are nothing if not clever), we can use the theory built up thus far to find the solution for negative time.

We define $g : [t_0 - h, t_0 + h] \times \overline{B(\xi_0, a)} \to \mathbb{R}^n$ where $g(t, \xi) = -f(2t_0 - t, \xi)$ (notice how as t increases from $t_0 - h$ to $t_0 + h$, the first input to f decreases from $t_0 + h$ to $t_0 - h$). We then find a solution on some $J' = [t_0, t_0 + b]$ (where b is the minimum value asserted in the statement of the theorem) to the IVP

$$\begin{aligned}y'(t) &= g(t, y(t)) \\ y(t_0) &= \xi_0\end{aligned}$$

By setting $x(t) = y(2t_0 - t)$ we obtain a solution on $[t_0 - b, t_0]$ as desired.

□

Now that we shown a few existence theorems, we should give uniqueness a shot as well. We will in fact give 2 proofs, related in essence but different in flavour.

Theorem 14.20 *Let D be an open subset of $\mathbb{R} \times \mathbb{R}^n$ with $(x_0, y_0) \in D$. Let $f : D \to \mathbb{R}^n$ be a function that is continuous in x and Lipschitz in y, with Lipschitz constant K. Then there exists some $a > 0$ such that the initial value problem*

$$\begin{aligned} y'(x) &= f(x, y(x)) \\ y(0) &= x_0 \end{aligned}$$

has a unique solutions on $(x_0 - a, x_0 + a)$.

Contraction Mapping. For the first proof, we will use Banach's contraction mapping theorem.

First we choose a rectangle $R' = [x_0 - A, x_0 + A] \times [y_0 - L, y_0 + L]$ (if $y_0 \in \mathbb{R}^n$ for $n > 1$, we do this for every coordinate) that is contained in D. Since f is continuous and R' is compact, it is bounded by some constant M. Let a be a positive real number that is less than $\min\{\frac{L}{M}, A, \frac{1}{K}\}$. Let $R = [x_0 - a, x_0 + a] \times [y_0 - L, y_0 + L]$.

Now we define a space $X := \{y \in C([x_0 - a, x_0 + a]; \mathbb{R}^n) : \|y - y_0\|_\infty \leq L\}$. Note that if $y \in X$, then the graph of y is contained in R' and since $R' \subset D$, in particular it is contained in D. Thus $f(x, y(x))$ makes sense for every $y \in X$. We would like to apply the Banach contraction mapping theorem to X (note it is complete as it is the closed subset of a complete metric space). Then we need a map $\Gamma : X \to X$ that is a contraction from which we will get a fixed point. Remembering the integral equation, what we want is to find a $y \in C([x_0 - a, x_0 + a]; \mathbb{R}^n)$ such that

$$y(x) = y_0 + \int_{x_0}^{x} f(s, y(s))ds$$

This tells us how to define Γ. In particular,

$$\Gamma(y)(x) = y_0 + \int_{x_0}^{x} f(s, y(s))ds$$

If we can find a fixed point of Γ, we will have found a solution to our initial value problem. In order to do so, all we need do is verify that the hypotheses of the Banach contraction mapping theorem hold.

First we want to know that Γ does indeed map X into itself (a priori, it is not obvious that the image of $y \in X$ under Γ is in X). Luckily for us this indeed the case since

$$\begin{aligned} |\Gamma(y)(x) - y_0| &= \left| \int_{x_0}^{x} f(s, y(s))ds \right| \\ &\leq \int_{x_0}^{x} |f(s, y(s))|ds \\ &\leq M|x - x_0| \\ &\leq Ma \\ &\leq L \end{aligned}$$

As this holds for all x, we get that

$$\|\Gamma(y) - y_0\|_\infty \leq L$$

The only thing that remains to be shown then, is that Γ is a contraction. What we will show is that $\|\Gamma(y) - \Gamma(z)\| \leq aK\|y - Z\|$. By our choice of a, we know that $aK < 1$ and so the claim will be satisfied.

$$\begin{aligned} |\Gamma(y)(x) - \Gamma(z)(x)| &\leq \int_{x_0}^{x} |f(s, y(s)) - f(s, z(s))|ds \\ &\leq \int_{x_0}^{x} K|y(s) - z(s)|ds \\ &\leq K \int_{x_0}^{x} \|y - z\|_\infty \\ &= K\|y - z\|_\infty (x - x_0) \\ &\leq aK\|y - z\|_\infty \end{aligned}$$

The second line follows from the fact that f is Lipschitz in the second component.

Thus we can safely apply the contraction mapping theorem and conclude the proof. Since the theorem tells us that the fixed point is unique, we know that our solution y is unique. □

Remark 14.21. Suppose y_1 is any continuous map. Then functions $\Gamma(y_1)$, $\Gamma^2(y_1) = \Gamma(\Gamma(y_1))$, $\Gamma^3(y_1) = \Gamma(\Gamma(\Gamma(y_1)))$,... are known as the Picard iterates.

Now we present a second proof of the same theorem using Picard iterates. Arguably this is the same thing, since if we were to follow through with the contraction mapping theorem, we would be applying the contraction, i.e. Γ, repeatedly to our initial point which is exactly what the Picard iterates are. However it's still useful to consider different proofs to see the different techniques they employ. Additionally, our solution will be defined on a slightly larger interval which is also nice.

Picard Iterates. Let R' be as before (in other words $R' = [x_0 - A, x_0 + A] \times [y_0 - L, y_0 + L]$ such that $R' \subset D$). This time we choose our a to be less that $\min\left\{\frac{L}{M}, A\right\}$ (where, as a reminder, M is a bound for f). As before, we will take $R = [x_0 - a, x_0 + a] \times [y_0 - L, y_0 + L]$.

Let $y_1 \equiv y_0$. For $n > 1$, we define $y_n = \Gamma(y_{n-1}) = \Gamma^{n-1}(y_1)$. We claim that the y_n converge uniformly to some $y \in C([x_0 - a, x_0 + a]; \mathbb{R}^n)$ where y is a solution to the integral equation (corresponding to the given IVP).

Perhaps, we should first show that the above makes sense. Which is to say, we need that $(x, y_n(x))$ is always in D so that we can evaluate f on it. But as shown in the previous proof, $\left\|\Gamma(y) - y_0\right\|_\infty \le L$ if $\left\|y - y_0\right\|_\infty \le L$. We know that $\left\|y_1 - y_0\right\|_\infty \le L$ (in fact it's 0!), so its graph is contained in R (which in turn is contained in D). By induction, this holds for all y_n.

Now we get to the actual meat: showing that the y_n converge. First we can immediately compute that

$$|y_2(x) - y_1(x)| \le M|x - x_0|$$

This a calculation we have done many times previously. Then we see that for $n \ge 2$

$$|y_{n+1}(x) - y_n(x)| \le \int_{x_0}^{x} |f(t, y_n(t)) - f(t, y_{n-1}(t)|dt \le K \int_{x_0}^{x} |y_n(t) - y_{n-1}(t)|dt$$

By induction then what we get is that

$$|y_{n+1}(x) - y_n(x)| \le K^{n-1} M \frac{|x - x_0|^n}{n!} \le K^{n-1} M \frac{a^n}{n!}$$

Since this holds for all $x \in [x_0 - a, x_0 + a]$ what we get is that

$$\left\|y_{n+1} - y_n\right\|_\infty \le \frac{M}{K} \frac{(Ka)^n}{n!}$$

Since the right hand side denotes the terms of a convergent series, we can conclude that the y_n form a Cauchy sequence and hence converge to some y (to be precise, in order to show Cauchy, we should show that $|y_n - y_m| < \epsilon$ eventually, as well let $n, m \to \infty$. However $\left\|y_n - y_m\right\|_\infty \le \left\|y_n - y_{n-1}\right\|_\infty + \left\|y_{n-1} - y_{n-2}\right\|_\infty + \cdots + \left\|y_{m+1} - y_m\right\|_\infty$ (assuming $n > m$). Each of these terms can be made arbitrarily small by the above estimate and so we are done).

The final thing we wish to show is that y does indeed solve the integral equation. We know that

$$y_n(x) = y_0 + \int_{x_0}^{x} f(t, y_{n-1}(t))dt$$

We can consider what happens as we let n tend to infinity on both sides. The left hand side goes to $y(x)$ by definition. As for the right hand side, what we see is that

$$\left|\int_{x_0}^{x} f(t, y_{n-1}(t))dt - f(t, y(t))dt\right| \le K \int_{x_0}^{x} |y_{n-1}(t) - y(t)|dt \le Ka\left\|y_{n-1} - y\right\|$$

Therefore the right hand side approaches

$$y_0 + \int_{x_0}^{x} f(t, y(t))dt$$

and we are done. □

14.3 Uniqueness of solutions

The reader has the right to complain because the previous theorem guarantees not only the existence of a solution but also its uniqueness. This was quietly swept under the rug for the second proof (and arguably swept under the rug of another theorem for the first one!). We remedy that over here. In fact to make up for it, we will provide 2 proofs of uniqueness.

Proposition 14.22 *The solution found using Picard iterates for Theorem 14.20 is unique.*

Proof 1. Suppose z is another solution to the IVP on $[x_0 - a, x_0 + a]$. Then again we have that the graph of z is contained in $R = [x_0 - a, x_0 + a] \times [y_0 - L, y_0 + L]$ (all the quantities are as defined in the proof). This is because

$$\begin{aligned}
|z(x) - y_0| &= |z(x) - z(x_0)| \\
&\leq \left\| z' \right\| |(x - x_0)| \\
&\leq Ma \\
&\leq L
\end{aligned}$$

where M, as before, is a bound for f and a is chosen to be smaller than $\frac{L}{M}$ ($\left\| z' \right\|$ means the operator norm of z').

We will show that $|y(x) - z(x)| = 0$ for every $x \in [x_0 - a, x_0 + a]$. Immediately, computing this quantity, what we find is that

$$|y(x) - z(x)| = \left| \int_{x_0}^{x} f(t, y(t)) - f(t, z(t)) dt \right| \leq K \int_{x_0}^{x} |y(t) - z(t)| dt \leq 2LK|x - x_0|$$

Then

$$\begin{aligned}
|y(x) - z(x)| &\leq K \int_{x_0}^{x} |y(t) - z(t)| dt \\
&\leq K \int_{x_0}^{x} 2LK(t - x_0) dt \\
&\leq \frac{2LK(x - x_0)^2}{2}
\end{aligned}$$

By induction what we find is that

$$|y(x) - z(x)| \leq \frac{2LK(x - x_0)^n}{n!}$$

for all n. Clearly the right hand side goes to 0 as $n \to \infty$ (the series converges to an exponential for example), thus $y(x) = z(x)$. ☐

Proof 2. For the second proof, we once again reach the same estimate as before

$$|y(x) - z(x)| \leq K \int_{x_0}^{x} |y(t) - z(t)| dt$$

Now we define $u(x) = |y(x) - z(x)|$ which we will show is 0. Note that by assumption u statisfies the following property

$$u(x) \leq K \int_{x_0}^{x} u(t) dt$$

This is known as Grönwall's inequality.

Define

$$U(x) = \int_{x_0}^{x} u(t) dt$$

Then we know that

$$U'(x) = u(x) \leq KU(x)$$

If we had an equality here, we could say that $U(x) = U(x_0)e^{K(x-x_0)}$. Grönwall tells us that the statement remains true if we replace the equalities with inequalities. Intuitively this is because $U(x_0) = U(x_0)e^{K(x_0-x_0)}$, so they 'start' at the same point but the right hand side grows faster (has larger gradient) so must always be greater than the left hand side. For more rigorous reasons, consider the following argument

$$\left(\frac{U(x)}{e^{K(x-x_0)}}\right)' = \frac{U'(x) - KU(x)}{e^{K(x-x_0)}} \leq 0$$

This means that the function is non-increasing so must reach its maximum at x_0 (we only consider the case for $x > x_0$ the other case is near identical). Therefore

$$\frac{U(x)}{e^{K(x-x_0)}} \leq \frac{U(x_0)}{e^{K(x_0-x_0)}}$$
$$U(x) \leq U(x_0)e^{K(x-x_0)}$$

Since $U(x_0) = 0$, we know that $U(x) \leq 0$ for all x. However, U is also the integral of a non-negative function, so $U(x)$ must always be non-negative as well. This must mean that U is 0 which in turn implies that u is 0. □

Grönwall's inequality exists in different levels of generality. A slightly more general case (than the one used above) is as follows.

Theorem 14.23 *If $U'(x) \leq K(x)U(x)$ where $K(x)$ is continuous and $U(x)$ is (of course) differentiable, then*

$$U(x) \leq U(x_0)\exp\left(\int_{x_0}^{x} K(t)dt\right)$$

Proof. Exercise for the student :(□

We have seen that under sufficiently good conditions, we can get unique solutions on intervals of the form $[x_0, x_0 + \delta_1]$ for some $\delta_1 > 0$. We can then run the proofs again so that our new solutions are on $[x_0 + \delta_1, x_0 + \delta_1 + \delta_2]$. Thus we obtain a sequence of δ_i. There are then 2 possibilities: either the δ_i form a convergent series or they don't. Since they are all positive, if the δ_i don't converge, then we get a solution defined on $[x_0, \infty)$. Otherwise the endpoint is simply the value that the series converges to. We have seen an example of the latter phenomena with the differential equation

$$x' = x^2$$

which we noted blew up in finite time.

Now that we have some knowledge of the existence and uniqueness of solutions, we can ask some question about these solutions. For example, how do the solutions depend on the initial condition x_0? What if we have a parameter in our differential equation, how do solutions vary as we vary the parameter? If we have solutions on some bounded interval, how and when can we extend the solutions to the maximal time interval and what is the behaviour at the end point of this interval, especially if the endpoint is finite?

We begin by first making our lives easier, i.e. removing questions that are equivalent. We claim that looking at the dependence on initial conditions is entirely equivalent to looking at dependence on parameters. For example suppose f is a function that depends on a parameter $\lambda \in \mathbb{R}^m$. We would then be attempting to solve the IVP

$$x'(t) = f(t, x(t), \lambda)$$
$$x(t_0) = x_0$$

Let x be a solution of this IVP. Then we can define

$$z(t) = \begin{pmatrix} x(t) \\ \lambda \end{pmatrix}$$

which satisfies the IVP

$$\begin{aligned} z'(t) &= g(t, z(t)) \\ z(t_0) &= w \end{aligned}$$

where

$$g(t, z(t)) = \begin{pmatrix} f(t, x(t), \lambda) \\ 0 \end{pmatrix}, w = \begin{pmatrix} x_0 \\ \lambda \end{pmatrix}$$

There also exists a more general uniqueness theorem called Osgood's Uniqueness Theorem.

Theorem 14.24 (Osgood's Uniqueness Theorem) *Suppose $D \subset \mathbb{R} \times \mathbb{R}^n$ is open and contains some (x_0, y_0). Assume that for all $(x, y_1), (x, y_2) \in D$ we have that*

$$|f(x, y_1) - f(x, y_2)| \le \varphi(|y_1 - y_1|)$$

where $\varphi : [0, \infty) \to [0, \infty)$ is continuous and $\varphi(0) = 0$. It also has the properties that for every $a > 0$, we have $\varphi(a) > 0$ and

$$\int_0^a \frac{1}{\varphi(u)} du = \infty$$

Then the initial value problem

$$\begin{aligned} y' &= f(x, y(x)) \\ y(x_0) &= y_0 \end{aligned}$$

has no more than one solution.

Before proving the theorem, a few remarks need to be made. First we will only prove the theorem for $n = 1$. With some more tools and techniques, the given proof can be modified to work in the general case, but we won't worry about that here.

Another important point is that we are not assuming f to be continuous, so we cause the Cauchy-Peano theorem to assert the existence of a solution. Indeed, we will not prove that a solution exists; only that if it does, it is unique. However, if we had $\varphi(x) = Kx$ for some $K > 0$, then f would be Lipschitz continuous and we reduce back to previous cases. Thus this theorem is certainly more general than what we seen thus far.

Proof. We will prove this using contradiction. Suppose we have two distinct solutions $y_1(x), y_2(x)$ on some interval (α, β) (containing x_0 of course). We define $z(x) = y_1(x) - y_2(x)$. This function satisfies the IVP

$$\begin{aligned} z'(x) &= f(x, y_1(x)) - f(x, y_2(x)) \\ z(x_0) &= 0 \end{aligned}$$

If $z(x) = 0$ everywhere then we are done. So suppose that there is some x_1 such that $z(x_1) \neq 0$. Then importantly

$$z'(x_1) \le \varphi(|z(x_1)|) < 2\varphi(|z(x_1)|)$$

We now proceed using a comparison argument which we split into cases. The first (and only) case we consider is when $x_1 > x_0$ and $y_1(x_1) > y_2(x_1)$ (the remaining cases are left as exercises for the student(s) :( ).

Let v be the solution to the IVP

$$\begin{aligned} v'(x) &= 2\varphi(v) \\ v(x_1) &= z(x_1) =: z_1 \end{aligned}$$

Note that $z_1 > 0$ by the assumption that $y_1(x_1) > y_2(x_1)$. We see that v and z are functions that agree on x_1 but $v' > z'$. Thus the graphs of v and z cannot intersect anywhere else. We will show this leads to a contradiction. In particular we will show that v is an increasing, positive function defined at least on $(-\infty, x_1]$ but since $z(x_0) = 0$, this means there must be another intersection for some $x < x_1$.

In order to verify the above statements for v, it would be nice to have a slightly more explicit formula for it. Since the differentiation equation for v is separable, we can at least get part of the way there. By separating the variables, we know that

$$\int_{v(x)}^{v(x_1)=z_1} \frac{1}{\varphi(v)} dv = \int_x^{x_1} 2dx = 2(x_1 - x)$$

Or at least this is what v should satisfy as the solution (consider substituting $v(x)$ on the right hand side). Thus we define the map v so that it makes the above statement true. This only makes sense, however, if for every x there exists exactly one v_x such that

$$\int_{v_x}^{z_1} \frac{1}{\varphi(v)} dv = 2(x_1 - x)$$

However this must be the case since $\frac{1}{\varphi}$ is positive so the integral is going to be monotone.

We claim that $v(x)$ is defined for all $x < x_1$. We first define

$$\Phi(y) = \int_y^{z_1} \frac{1}{\varphi(u)} du$$

Then in particular $v(x)$ is the implicit solution to the equation

$$\Phi(v(x)) = 2(x_1 - x)$$

By assumption, $\Phi(y) \to \infty$ as $y \to 0$ (recall this was one of the properties of φ). Additionally $\Phi(z_1) = 0$ so $\Phi : (0, z_1] \to [0, \infty)$. Since $\Phi' = -\frac{1}{\varphi}$, we can conclude that Φ is decreasing. This mean that Φ is in particular invertible, allowing us to write

$$v(x) = \Phi^{-1}(2(x_1 - x))$$

The domain of Φ^{-1} is $[0, \infty)$ thus the above formula is define whenever $x_1 - x > 0$ or in other words when $x < x_0$.

Clearly v is increasing (there are 2 ways of seeing this: for one it is the composition of two decreasing maps and secondly, we can look at the IVP defining it and use the fact that φ is positive on non-zero values). Additionally, since Φ^{-1} takes values in $(0, z_1]$, v must also be positive. As discussed previously, this implies that the graphs of v and z intersect (at least) twice: once at z_1 and once prior, leading to the desired contradiction. □

Remark 14.25. There is a simpler proof if Φ were concave up which would occur if ϕ were non-decreasing. There is a standard way of turning any function into a non-decreasing function, define

$$\tilde{\varphi}(u) = \sup_{\tilde{u} \in [0,u]} \varphi(\tilde{u})$$

Unfortunately, it is not true that

$$\int_0^a \frac{1}{\varphi(u)} du = \infty \Rightarrow \int_0^a \frac{1}{\tilde{\varphi}(u)} du = \infty$$

Remark 14.26. If $\varphi : [0, \infty) \to [0, \infty)$ such that $\varphi(0) = 0$, $\varphi(a) > 0$ for $a > 0$ and $\varphi'(0)$ exists, then

$$\int_0^a \frac{1}{\varphi(u)} du = \infty$$

for all $a > 0$

Theorem 14.27 *Let (X,d) be a complete metric space. Suppose $F : X \times I \to X$ (where I is any interval) is such that $F(\cdot,\lambda)$ is a uniform contraction. In other words, there exists some $0 \le q < 1$ such that for every $\lambda \in I$ and every $x, y \in X$, we have that*

$$d(F(x,\lambda) - F(y,\lambda)) \le q d(x,y)$$

Then there exists a unique fixed point x_λ^*. This defines a map $\lambda \mapsto x_\lambda^*$ that satisfies

$$d(x^*(\lambda), x^*(\lambda')) \le \frac{1}{1-q} d(F(x^*(\lambda),\lambda), F(x^*(\lambda),\lambda'))$$

If F is continuous in λ then the map x^* is continuous (in λ) as well.

Remark 14.28. The word continuous in the final sentence can be replaced with Lipschitz, differentiable, C^k, etc.

Proof. The first statement for the bound on $d(x^*(\lambda), x^*(\lambda'))$ can be seen by

$$\begin{aligned} d(x^*(\lambda), x^*(\lambda')) &= d(F(x^*(\lambda),\lambda), F(x^*(\lambda'),\lambda')) \\ &\le d(F(x^*(\lambda),\lambda), F(x^*(\lambda),\lambda')) + d(F(x^*(\lambda),\lambda'), F(x^*(\lambda'),\lambda')) \\ &\le d(F(x^*(\lambda),\lambda), F(x^*(\lambda),\lambda')) + q d(x^*(\lambda), x^*(\lambda')) \end{aligned}$$

From this it is clear that continuity of F (in λ) implies continuity of x^*. It is not clear how this would generalise for smooth or C^k F. So we present another proof that will give these to us readily.

Suppose F is C^1. We seek to solve for x such that

$$x = F(x,v)$$

More precisely we know that for every v there is a solution and we wish to decide how these solutions depend on v. Taking everything to one side and taking derivatives what we find is that

$$I - D_x F = 0$$

where I is the identity. Since $\|D_x F\| \le q < 1$ what we find is that this is invertible (see following lemma). Therefore we can use the inverse function theorem to conclude that the dependence on solutions is smooth/C^K/etc. □

Lemma 14.29 *If $A \in \mathbb{R}^{n\times n}$ such that $\|A\| < 1$ then $I - A$ is invertible.*

Proof. The inverse of $I - A$ is given by

$$\sum_{k=0}^{\infty} A^k$$

(note similarity to geometric series). The series converges absolutely since

$$\sum_{k=0}^{\infty} \left\| A^k \right\| \le \sum_{k=0}^{\infty} \|A\|^k$$

Then we see that

$$\begin{aligned} (I-A)\sum_{k=0}^{\infty} A^k &= \lim_{N\to\infty} (I-A)\sum_{k=0}^{N} A^k \\ &= \lim_{N\to\infty} \sum_{k=0}^{N} A^k - \sum_{k=0}^{N} A^{k+1} \\ &= \lim_{N\to\infty} I - A^{N+1} \end{aligned}$$

Since A^N converges to 0 (check norms), we get the desired statement. □

Remark 14.30. A remark that is entirely unrelated to ODEs but also so delightful that I cannot help but mention it here. We've seen that plugging matrix into e^x and $\frac{1}{1+x}$ kind of makes sense (admittedly the latter only working for some matrices). In both cases we used the fact that these functions can be approximated via polynomials. Thus as you might imagine, any time we can approximate a function via polynomials, we can consider what happens if we input a matrix! The even more amazing fact is that any continuous function on a compact interval can be approximated via polynomials. So in fact we can apply any continuous function to a matrix. This is what leads to the so-called functional calculus.

Definition 14.31 (Locally Lipschitz)**.** A map $f: A \to \mathbb{R}^n$ is said to be locally Lipschitz, if for every compact set $K \subset A$ there exists a constant L (which may depend on K) such that $|f(x) - f(y)| \leq L|x-y|$ for all $x, y \in K$.

We will finish with a global existence and uniqueness theorem. Essentially, what we want to say is that if we have a solution on the maximal time interval then the solution blows up (in norm) or the solution approaches the boundary of the domain (or both).

Theorem 14.32 *Suppose $D \subset \mathbb{R}^n$ is open and connected and $f: D \to \mathbb{R}^n$ is a locally Lipschitz vector field. Let v be some vector in D. Then there exists a unique maximal interval of existence $I_{max} = (\underline{T}, \overline{T}) \ni 0$ such that*

1. *The IVP*

$$x' = f(x)$$
$$x(0) = v$$

   has a unique solution on I_{max} and

2. *if $\overline{T} < \infty$ then*

$$\lim_{t \to \overline{T}} |x(t)| + \frac{1}{d(x(t), \partial D)}$$

Proof. We already know that the solution exists and is unique so we only need to verify the second statement. We do so by showing the contrapositive. In particular we will show that if $T < \overline{T}$ such that $\lim_{t\to T} |x(t)| \neq \infty$ and $\lim_{t \to T} d(x(t), \partial D) \neq 0$ then there is a solution that goes past T.

Let $T > 0$ be arbitrary and suppose the two limit conditions from the previous paragraphs hold (that is $|x(t)|$ doesn't tend to infinity and $x(t)$ always remains some distance from the boundary). Then there exists a sequence $(t_j)_{j \in \mathbb{N}}$ that converges to T such that $M := \sup_{j \in \mathbb{N}} |x(t_j)|$ is finite and $S := \inf_{j \in \mathbb{N}} d(x(t_j), \partial D)$ is positive (or in other words is not 0). Define

$$K = \{z \in D : |z| \leq M, d(z, \partial D) \geq S\}$$

K is closed and bounded, hence it is compact (note that the fact that K is closed might not be as apparent as it seems. In particular $K = D \cap \overline{B(0,M)} \cap \{z \in \mathbb{R}^n : d(z, \partial D) \geq S\}$. The final set is certainly closed however since D is open, it is not clear that the intersection should also be closed. Nevertheless it is true that K is closed because $D \cap \{z \in \mathbb{R}^n : d(z, \partial D) \geq S\} = \{z \in \mathbb{R}^n : d(z, D^c) \geq S\}$ which *is* obviously closed being the preimage of a closed set under a continuous function).

On K, then we have a Lipschitz constant L and hence by theory we have built up so far (see Theorem 14.20) we know there is some $\tau > 0$ such that for every $w \in K$ we have a unique solution to the IVP

$$y' = f(y)$$
$$y(0) = w$$

on $[-\tau, \tau]$ (to be fair we started with an open subset of $\mathbb{R}^{n+1}$ in that proof whereas we have a compact set here. However, we only needed the fact that the set contained a rectangle. Given the definition of K, we can see that this should be true).

Let y_j be the solutions for $w = x(t_j)$ for the above IVP. Then $y_j(t - t_j)$ is defined on $[\tau - t_j, \tau + t_j]$ which agrees with $x(t)$ at t_j (note that since K contains all the t_j by its definition, this τ holds for all t_j). Since x is maximal and unique, we conclude that x is defined at least until $t_j + \tau$. Now suppose we pick some t_j such that $T - t_j < \tau$ (we can definitely do so since t_j converge to T). But this means that $x(t)$ is defined at least till $t_j + \tau > T$. This holds for all T so $x(t)$ is defined on $[0, \infty)$. □

15 Continuous Dependence

We know now that under certain conditions solutions exist and are unique and we even have that solutions vary continuously as the initial conditions (or parameters) vary continuously. This means, for example then, if we have a sequence v_k in $\mathbb{R}^n$ that converge to some v then the solution to

$$\begin{aligned} x' &= f(t,x) \\ x(0) &= v \end{aligned}$$

can be found by taking the appropriate limits with the v_k. In particular, if we use x_k to denote the solutions to

$$\begin{aligned} x' &= f(t,x) \\ x(0) &= v_k \end{aligned}$$

Then the solution to the original IVP can be found by $x(t) = \lim_{k\to\infty} x_k(t)$ (again, assuming suitable conditions for f). This is easier to express using flow notation as we can equivalently express the above as $\lim_{k\to\infty} \Phi_t(v_k) = \Phi_t(v)$ for all t. (see Section 13 for definition of the flow map). Although this is lovely, it would be even more lovely if we could have an estimate for how fast the convergence is. To get there, we first need a generalised version of Grönwall's lemma.

Theorem 15.1 (Grönwall's Lemma (Generalised)) *Suppose we have continuous functions $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}^+$. Suppose $y : [a,b] \to \mathbb{R}$ is continuous and satisfies*

$$y(t) \le f(t) + \int_a^t g(s)y(s)ds$$

for all $t \in [a,b]$. Then we have that

$$y(t) \le f(t) + \int_a^t f(s)g(s)\exp\left(\int_s^t g(u)du\right)ds$$

In particular if $f(t) \equiv k$ for some constant k, then

$$y(t) \le k\exp\left(\int_a^t g(u)du\right)$$

First let us verify that the statement works for $f(t) \equiv k$. In other words we wish to show that

$$k + k\int_a^t g(s)\exp\left(\int_s^t g(u)du\right)ds \le k\exp\left(\int_a^t g(u)du\right)$$

We see that the above expressions are equal at $t = a$. Thus if we had that the derivative of the left hand side was always less than the right, we would have the desired inequality.

First we rewrite the left hand side as

$$\begin{aligned} k + k\int_a^t g(s)\exp\left(\int_s^t g(u)du\right)ds &= k + k\int_a^t g(s)\exp\left(\int_a^t g(u)du - \int_a^s g(u)du\right)ds \\ &= k + k\exp\left(\int_a^t g(u)du\right)\cdot\int_a^t g(s)\exp\left(-\int_a^s g(u)du\right)ds \end{aligned}$$

Substituting this into the the desired inequality and multiplying both sides by $\frac{1}{k}\exp(-\int_a^t g(u)du)$, we see that the inequality is equivalent to

$$\exp\left(-\int_a^t g(u)du\right) + \int_a^t g(s)\exp\left(-\int_a^s g(u)du\right)ds \le 1$$

Thus we will show that this inequality is true. Once again, it is clear that we have agreement on $t = a$, so we only need argue that the inequality remains with derivatives. Taking derivatives on the left hand side (with respect to t) we get

$$\exp\left(-\int_a^t g(u)du\right)\cdot -g(t) + g(t)\exp\left(-\int_a^t g(u)du\right) = 0$$

Therefore the expression on the left is actually constant! Thus rather than having inequality we have equality.

There is in fact another we can verify equality that is a bit more straightforward and in fact employs a simple technique for computing such double integrals that we will use soon so is worth mentioning. Recall we are trying to evaluate

$$k + k\int_a^t g(s)\exp\left(\int_s^t g(u)du\right)ds$$

We define

$$M(s) = \int_s^t g(u)du = -\int_t^s g(u)du$$

Then

$$\begin{aligned}
k + k\int_a^t g(s)\exp\left(\int_s^t g(u)du\right)ds &= k + k\int_a^t -M'(s)\exp(M(s))ds \\
&= k - k(\exp(M(s)|_a^t) \\
&= k - k(\exp(M(t)) - \exp(M(a))) \\
&= k\exp(M(a)) \\
&= k\exp\left(\int_a^t g(u)du\right)
\end{aligned}$$

Having verified the theorem in some ways, we can try proving it.

Proof 1. For the first proof we will iterate the the assumed inequality. We will see that doing so gives us the Taylor series of an exponential with an additional term that goes to 0 as we continually iterate the inequality.

First note that without loss of generality we can assume $|y(t)| \leq 1$ (y is continuous on a compact set so $\sup_{t\in[a,b]}|y(t)|$ is finite. Dividing by this quantity, we see that the assumed inequality is maintained). Additionally we will only show the case for when f is a constant (the general case is left as an exercise).

By assumption we know that

$$y(t) \leq k + \int_a^t g(s)y(s)ds$$

Applying this inequality to the integral what we conclude is that

$$y(t) \leq k + \int_a^t g(s)y(s)ds \leq k + \int_a^t g(t_1)\left[k + \int_a^{t_1} g(s)y(s)ds\right]dt_1$$

Expanding the right hand side, what we get is

$$k + \int_a^t g(t_1)\left[k + \int_a^{t_1} g(s)y(s)ds\right]dt_1 = k + k\int_a^t g(t_1)dt_1 + \int_a^t g(t_1)\int_a^{t_1} g(s)y(s)ds\, dt_1$$

Unfortunately we need to apply the inequality again to get something useful. Doing so and expanding it like above, what we find is that

$$y(t) \leq k + k\int_a^t g(t_1)dt_1 + k\int_a^t g(t_1)\int_a^{t_1} g(t_2)dt_2\, dt_1 + k\int_a^t g(t_1)\int_a^{t_1} g(t_2)\int_a^{t_2} g(t_3)y(t_3)dt_3\, dt_2\, dt_1$$

Although this is quite a mess, we won't worry about all of it (for now). Instead just focus on the double integral term. As a first step we will name it (very often a useful first step)

$$I = \int_a^t g(t_1)\int_a^{t_1} g(t_2)dt_2\, dt_1 = \int_a^t M'(t_1)M(t_1)dt_1$$

where

$$M(t_1) = \int_a^{t_1} g(s)ds$$

Therefore

$$I = \frac{1}{2}[M(t)]^2 - \frac{1}{2}[M(a)]^2$$

However, $M(a) = 0$ so

$$I = \frac{1}{2}[M(t)]^2 = \frac{1}{2}\left(\int_a^t g(s)ds\right)^2$$

With a bit of working out or a bit of faith, we conclude that repeatedly applying the inequality to the integrals gets us

$$y(t) \le k + k\int_a^t g(t_1)dt_1 + \frac{k}{2!}\left(\int_a^t g(t_1)dt_1\right)^2 + \cdots + \frac{k}{n!}\left(\int_a^t g(t_1)dt_1\right)^n + \int_a^t g(t_1)\cdots\int_a^{t_n} g(t_{n+1})y(t_{n+1})dt_{n+1}\ldots dt_1$$

For completeness, we can verify this holds for the triple integral term which would appear after we apply the inequality to the integral again. In other words we are trying to evaluate

$$J = \int_a^t g(t_1)\int_a^{t_1} g(t_2)\int_a^{t_2} g(t_3)dt_3\, dt_2\, dt_1$$

Using our previous notation and result, we see that

$$J = \int_a^t M'(t_1)\cdot\frac{1}{2}[M(t_1)]^2 dt_1$$

Therefore

$$J = \frac{1}{3!}[M(t)]^3$$

as was claimed. Hopefully we can convince ourselves now that the above summation is correct. We can see the desired exponential term appearing (as a Taylor series). Thus if the claim is to hold true then we should find that the final term goes to 0 as we repeatedly apply the inequality. Now we verify that this happens.

Since g is always positive what we find is that

$$\left|\int_a^t g(t_1)\cdots\int_a^{t_n} g(t_{n+1})y(t_{n+1})dt_{n+1}\ldots dt_1\right| \le \int_a^t g(t_1)\cdots\int_a^{t_n} g(t_{n+1})|y(t_{n+1})|dt_{n+1}\ldots dt_1$$

We assumed that $|y(t)| \le 1$ for all $t \in [a,b]$. Therefore the right hand side above is less than or equal to

$$\int_a^t g(t_1)\cdots\int_a^{t_n} g(t_{n+1})dt_{n+1}\ldots dt_1$$

However we already know that this is equal to

$$\frac{1}{n!}\left(\int_a^t g(s)ds\right)^n$$

Clearly this goes to 0 as $n \to \infty$ (it is a term of a convergent infinite series for example). Therefore we conclude that

$$y(t) \le k\exp\left(\int_a^t g(s)ds\right)$$

□

As usual there is a second proof of the theorem. This one, a bit more indirect.

Proof 2. We define
$$z(t)=\int_a^t g(s)y(s)ds$$
By assumption we have that
$$y(t)\le f(t)+\int_a^t g(s)y(s)ds$$
By multiplying both sides by $g(t)$ (remember g is positive) and rearranging what we find is that
$$g(t)y(t)-g(t)\int_a^t g(s)y(s)ds\le g(t)f(t)$$
Therefore
$$z'(t)-g(t)z(t)\le f(t)g(t)$$
Seeing the left hand side one is reminded of linear exact differential equations and integrating factors (see Section 12). In particular, we can write the left hand side as the derivative of a product if we multiply both sides by
$$\exp\left(-\int_a^t g(u)du\right)$$
(note the inequality is maintained since the exp is always positive). The inequality now becomes
$$z'(t)\exp\left(-\int_a^t g(u)du\right)-z(t)g(t)\exp\left(-\int_a^t g(u)du\right)\le f(t)g(t)\exp\left(-\int_a^t g(u)du\right)$$
Although the left hand side looks like something of a mess, we know (by construction) that it is equal to $w'(t)$ where
$$w(t)=z(t)\exp\left(-\int_a^t g(u)du\right)$$
Hence we get
$$\begin{aligned}\int_a^t w'(s)ds&\le\int_a^t f(s)g(s)\exp\left(-\int_a^s g(u)du\right)ds\\ z(t)\exp\left(-\int_a^t g(u)du\right)&\le\int_a^t f(s)g(s)\exp\left(-\int_a^s g(u)du\right)ds\\ z(t)&\le\int_a^t f(s)g(s)\exp\left(\int_s^t g(u)du\right)ds\\ f(t)+\int_a^t g(s)y(s)ds&\le f(t)+\int_a^t f(s)g(s)\exp\left(\int_s^t g(u)du\right)ds\end{aligned}$$
where for the second line we use the fact that $w(a)=0$ (which follows from the fact that $z(a)=0$). Since the left hand side is greater or equal to $y(t)$ by assumption, we are done. □

Now that we have Grönwall's Lemma, we can use it to to provide estimates for how much solutions can differ, if we only change the initial conditions slightly. In particular, we have the following theorem.

Theorem 15.2 *Let $x_1:[a,b]\to\mathbb{R}^n$ and $x_2:[a,b]\to\mathbb{R}^n$ be differential functions such that $|x_1(a)-x_2(a)|\le\delta$. Let $f:[a,b]\times\mathbb{R}^n\to\mathbb{R}^n$ be function that is Lipschitz in the second variable with Lipschitz constant L. Suppose*
$$\begin{aligned}|x_1'(t)-f(t,x_1(t))|&\le\epsilon_1\\ |x_2'(t)-f(t,x_2(t))|&\le\epsilon_2\end{aligned}$$
Then
$$|x_1(t)-x_2(t)|\le\delta e^{L(t-a)}+(\epsilon_1+\epsilon_2)\frac{e^{L(t-a)}-1}{L}$$

Remark 15.3. Suppose $\epsilon_1 = \epsilon_2 = 0$. Then x_1 and x_2 satisfy the same differential equation but with slightly different initial conditions. In this case the conclusion of the theorem is that $|x_1(t) - x_2(t)| \le \delta e^{L(t-a)}$, which is to say that the solutions can differ at most by some exponential term. If we also have $\delta = 0$, we get yet another proof of uniqueness.

Proof. Let $\epsilon = \epsilon_1 + \epsilon_2$. We define $g(t) = x_1(t) - x_2(t)$. We want to show that g is bounded by some kind of exponential term and as we said we want to use Grönwall. For this we need a bound on g. First we will find an estimate for g' (as you can imagine this will be easier since we have some strong conditions on f which approximates x_1' and x_2' fairly well) and then use this to find an estimate for g.

$$\begin{aligned}
|g'(t)| &= |x_1'(t) - x_2'(t)| \\
&= |(x_1'(t) - x_2'(t)) - f(t, x_1(t)) + f(t, x_1(t)) - f(t, x_2(t)) + f(t, x_2(t))| \\
&\le |f(t, x_1(t)) - f(t, x_2(t))| + |x_1'(t) - f(t, x_1(t))| + |x_2'(t) - f(t, x_2(t))| \\
&\le L|g(t)| + \epsilon
\end{aligned}$$

Now that we have an estimate for g' we can try to find an estimate for g as well. In particular we get

$$\begin{aligned}
|g(t)| &= \left| g(a) + \int_a^t g'(s) ds \right| \\
&\le |g(a)| + \int_a^t |g'(s)| ds \\
&\le \delta + \int_a^t L|g(s)| + \epsilon ds \\
&= \delta + \epsilon(t-a) + \int_a^t L|g(s)| ds
\end{aligned}$$

Using Grönwall then what we conclude is that

$$|g(t)| \le \delta + \epsilon(t-a) + \int_a^t L(\delta + \epsilon(s-a)) e^{L(t-s)} ds$$

Computing the right hand side, we find it to be equal to

$$\delta e^{L(t-a)} + \frac{\epsilon}{L}(e^{L(t-a)} - 1)$$

giving us the stated inequality. □

We now make a statement about solving (nonautomous) linear systems that could have been stated earlier but is here because we will use it shortly.

Corollary 15.3.1 *Let $(t_0, x_0) \in I \times \mathbb{R}^n$ be arbitrary, where I is some interval in $\mathbb{R}$. Let $A(t)$ be a continuous family of $n \times n$ matrices. Then the IVP*

$$\begin{aligned}
x'(t) &= A(t)x \\
x(0) &= x_0
\end{aligned}$$

has a unique solution on all of I. Moreover this solution is given by

$$x(t) = e^{\int_{t_0}^t A(s) ds} x_0$$

Proof. We already know that the solution exists and is unique from everything we've done. We can check that the solution is as given by verifying that it satisfies the ODE. Indeed we see that

$$x'(t) = A(t) e^{\int_{t_0}^t A(s) ds} x_0 = A(t)x$$

and $x(t_0) = x_0$. By uniqness, this is the solution to the IVP. □

16 Flow of Dynamical Systems

Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is C^1. Then we for every $x \in \mathbb{R}^n$ there is a unique solution satisfying the IVP

$$\begin{aligned} y' &= f(y) \\ y(0) &= x \end{aligned}$$

Suppose also that the solution is defined for all $t \in \mathbb{R}$. Then we define $\Phi_t(x) = y(t)$ (where y is the (unique) solution to the IVP). In effect Φ_t maps initial points x to where the differential equation sends them at time t. As mentioned in Section 13, the flow map satisfies the semigroup property: $\Phi_{s+t} = \Phi_s \circ \Phi_t$ (note this property does *not* hold if we do not have uniqueness of solutions).

Sometimes we alternatively use the notation $\Phi(t, x)$. This is especially useful when we wish to take the derivative with respect to time for example. In fact let us do exactly this to try and calculate the so-called time derivative $\frac{\partial \Phi}{\partial t}$. This means that we hold the initial value x constant and vary t. Clearly this must be the solution $y(t)$. Therefore

$$\frac{\partial \Phi}{\partial t}(t, x) = f(\Phi(t, x))$$

Additionally, it is clear by definition that $\Phi_0(x) = \Phi(0, x) = x$, so Φ_0 is the identity. From these two facts we can conclude that Φ_t is a bijection for all t with inverse Φ_{-t}. The injectivity of Φ_t is exactly equivalent to a previous statement we have made of how the uniqueness of solutions implies that the solutions can never intersect. The surjectivity is a confirmation of a long held suspicion of ours that the general solution for $y' = f(y)$ where $f : \mathbb{R}^n \to \mathbb{R}^n$ is an $n-$dimensional space.

Now that we have considered the time derivative, we should also consider its sibling the space derivative $\frac{\partial \Phi}{\partial x}$. First note we know this exists and is continuous by Theorem 14.27 (or more precisely by the remark following it). As you can imagine however, computing it will be harder.

Consider our standard IVP

$$\begin{aligned} x'(t) &= f(x(t)) \\ x(0) &= x_0 \end{aligned}$$

for $t \in J$ where J is a closed interval containing 0 and f is C^1. For $t \in J$, we define $A(t) = Df(x(t))$ (this is the Jacobian of f at $x(t)$). Since f is C^1, we know that A is continuous function. We then define the *variational equation along the solution* $x(t)$ to be the unique solution $u(t)$ to

$$\begin{aligned} u' &= A(t)u \\ u(0) &= u_0 \end{aligned}$$

(recall we know that unique solutions exist for every u_0 and are defined on all of J by Corollary 15.3.1). Note that this is a (admittedly non-autonomous) linear system of equations, hence the linearity principle still holds. What we will show is that if u_0 is small the map $t \mapsto x(t) + u(t)$ (with $x(t)$ as above) is a good approximation to the solution of

$$\begin{aligned} x' &= f(x) \\ x(0) &= x_0 + u_0 \end{aligned}$$

We formalise this in the following proposition.

Proposition 16.1 *Let $D \subset \mathbb{R}^n$ and $f : D \to \mathbb{R}^n$ be C^1. Let J be a closed interval containing 0 and $x(t)$ is the solution to $x' = f(x)$ with $x(0) = x_0$. Let $u(t)$ be the solution to $u' = Df(x(t))u$ satisfying $u(0) = \xi$ (in other words u is the variation equation along $x(t)$) and let $y(t)$ be the solution to $x' = f(x)$ satisfying $y(0) = x_0 + \xi$. Then*

$$\lim_{\xi \to 0} \frac{|y(t) - x(t) - u(t)|}{|\xi|}$$

converges uniformly to 0.

Remark 16.2. Uniform convergence in this context means that for every $\epsilon > 0$ there exists some $\delta > 0$ such that if $|\xi| \le \delta$ then $|y(t) - x(t) - u(t)| \le \epsilon|\xi|$ for every $t \in J$.

Proof. As usual we go to integral equations. We have

$$x(t) = x_0 + \int_0^t f(x(s))ds$$
$$y(t) = x_0 + \xi + \int_0^t f(y(s))ds$$
$$u(t) = \xi + \int_0^t Df(x(s))u(s)ds$$

Then

$$|y(t) - x(t) - u(t)| \le \int_0^t |f(y(s)) - f(x(s)) - Df(x(s))u(s)|ds \tag{16.1}$$

Let us denote the left hand side as $g(t)$. Finding the Taylor expansion of f centered at x what we find is that

$$f(y) = f(x) + Df(x)(y-x) + R(y-x)$$

where $R(y-x)$ is the remainder term such that

$$\lim_{y \to x} \frac{R(y-x)}{|y-x|} = 0$$

Substituting this expansion of f into the above inequality, we get

$$\begin{aligned} g(t) &\le \int_0^t \big|f(x(s)) + Df(x(s))(y(s)-x(s)) + R(y(s)-x(s)) - f(x(s)) - Df(x(s))u(s)\big|\,ds \\ &\le \int_0^t |Df(x(s))(y(s)-x(s)-u(s))|ds + \int_0^t |R(y(s)-x(s))|ds \end{aligned}$$

Since J is a closed interval it is in particular compact hence $|Df(x(s))|_{s\in J}$ achieves its maximum. Let N be this maximum. Then N serves as a Lipschitz constant for f. Let $\epsilon > 0$ be given. By uniform convergence of $\frac{R(y-x)}{|y-x|}$ to 0, we know there is some $\delta_1 > 0$ so that $|y-x| < \delta_1$ then $|R(y-x)| \le \epsilon|y-x|$. But recall that y and x are solutions to the same differential equation with the differing initial condition. We already have an estimate for how far these can differ, by Theorem 15.2. We know f is C^1 so it is locally Lipschitz. Thus for every compact set C, we can find a constant K such that f restricted to C is Lipschitz with Lipschitz constant K. There is a compact set containing the images of x and y (we can in particular take the union of their images). Let K be the Lipschitz constant for this compact set. Then we know by the aforementioned theorem that

$$|y(s) - x(s)| \le |\xi|e^{Ks}$$

for every $s \in J$. Since J is compact the right hand attains some maximum. Thus by choosing an appropriate $\delta > 0$ we can ensure that if $|\xi| < \delta$ then $|\xi|e^{Ks} < \delta_1$. Thus if $|\xi| < \delta$, then we know that $|R(y(s)-x(s)| \le \epsilon|y(s)-x(s)| \le \epsilon|\xi|e^{Ks}$ for every $s \in J$. Under this condition, what we find then is that

$$\begin{aligned} g(t) &\le \int_0^t |Df(x(s))(y(s)-x(s)-u(s))|ds + \int_0^t |R(y(s)-x(s))|ds \\ &\le \int_0^t |Df(x(s))|g(s)ds + \int_0^t \epsilon|\xi|e^{Ks}ds \\ &\le \int_0^t Ng(s)ds + \epsilon|\xi|\int_J e^{Ks}ds \end{aligned}$$

Note that the integral in the second term is simply a constant which we will denote c. Thus using Grönwall (specifically the case when f is a constant function), we conclude that

$$g(t) \le c\epsilon|\xi|e^{Nt}$$

Thus

$$\lim_{\xi\to 0}\frac{g(t)}{|\xi|}$$

goes to 0 uniformly since ϵ was chosen arbitrarily. □

Why is this proposition important? Well, as we've said it is often impossible to solve differential equations explicitly. However, perhaps we can find $x(t)$ for some special x_0 (for example, we might be able to find equilibria, which would would correspond to the zeroes of f). This means that we can approximate solutions near these special x_0 to get at least get some idea of the behaviour. In fact we can compute $\frac{\partial\Phi}{\partial x}$ a bit more explicitly now. Then $y(t)$ and $x(t)$ in the above proposition are simply $\Phi(t, x_0+\xi)$ and $\Phi(t, x_0)$. We think of $u(t)$ as a function of 2 variables: namely we define $u(t,\xi)$ to be the solution to the variational equation along $x(t)$ (or $\Phi(t, x_0)$) satisfying $u(0,\xi) = \xi$. The linearity principle ensure that u is linear in ξ. The definition of the differential of f is the (unique) linear map A that satisfies

$$\lim_{h\to 0}\frac{|f(x-h)-f(x)-Ah|}{|h|}=0$$

Therefore the map $\xi \mapsto u(t,\xi)$ is the partial derivative of $\Phi(t,x)$ with respect to x, which we may also denote $D\Phi_t(x)$. In fact since we have an explicit solution for u, we can write

$$D\Phi_t(x_0) = e^{\int_0^t Df(x(s))ds}x_0$$

We can then verify that $D\Phi_t(x)$ satisfies the following IVP

$$\begin{aligned}&\frac{\partial}{\partial t}(D\Phi_t(x)) = Df(\Phi_t(x))D\Phi_t(x)\\ &D\Phi_0(x) = I\end{aligned}$$

where I is the identity map. This matches the IVP for u as well (we use the fact that $A(t) = Df(x(t)) = Df(\Phi_t(x_0))$ and $u(0, x_0) = x_0$).

To recap, we have managed to find some formulae and descriptions of the variation equation. The reason we wanted to solve the variation equation is because it lets us approximate solutions around known solutions. The question then is, have we actually done this? We can find $u(t)$ by solving an integral or by solving an initial value problem, neither of which is particularly easy in general. And indeed, what we've done is converted one difficult problem into another (hurrah.). However there are some important cases where we can solve for things and are hence worth exploring.

Suppose the known solution $x(t)$ is an equlibrium point. In other words $x(t) \equiv a$ for some a. Then $A(t) = Df(a)$ for all t. Hence what we find is that

$$u(t) = D\Phi_t(a) = e^{tDf(a)}$$

Therefore we conclude that in a neighbourhood around equilibria, the flow is approximated by a linear system.

16.1 Classification of Flow

Much like we did with linear systems in the plane, we want to classify all (possibly non-linear) systems in any number of dimensions. This as daunting a task as it sounds. At the very least, we first need to start with how this classification will work, namely when can we say that two flows are 'the same'? This leads us to the notion of similarity.

We will say that two flows Φ and Ψ are similar (or more precisely conjugate) when there is a bijection h such that

$$h^{-1}\circ\Phi\circ h = \Psi$$

Remark 16.3. Note how similar this is to similarity of matrices: we say two matrices A, B are similar if there is an invertible T such that $T^{-1}AT = B$.

Note how we have not made any assumption about h being continuous/differentiable/smooth/etc. Such assumptions can vary based on the field of study. For our purposes, we will usually take h to be a homeomorphism (although if it is a diffeomorphism, that would be even more lovely!).

Now suppose x is a solution to the non-linear system

$$\begin{aligned} x' &= f(x) \\ x(0) &= v \end{aligned}$$

and z is its linearization. In other words, z is the solution to

$$\begin{aligned} z'(t) &= Df(x(t))z \\ z(0) &= w \end{aligned}$$

(where w is taken to be small). As we've discussed, $x+z$ is an approximation solution to $x' = f(x)$ with $x(0) = v + w$ (this is exactly Proposition 16.1). Thus we would like to say that x is conjugate to $x+z$. Unfortunately in general this is not true, even in the simplest case where we approximate near a steady state. Consider the following example.

16.1.1 Example 1

Suppose we are given the differential equation

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix}$$

Clearly, the only equilibrium point is the origin and

$$Df(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Thus the linearised system has constant solutions. And indeed, if we were to look in a small neighbourhood near the origin the solutions (on a small time interval) do look roughly constant (see here). However, and this is the key, the constant solution is *not* conjugate to the non-linear system. This is essentially for the exact reason you would suspect as well: although the solutions look constant for small intervals, for large enough intervals, this is no longer the case. It might be a useful exercise to prove to yourself that constant systems cannot be conjugate to non-constant ones, before I do that right now.

Quite simply, the key is that if the solutions to a system are constant then the flow map Φ_t is simply the identity for all t. Thus $h \circ \Phi_t \circ h^{-1}$ must be the identity as well.

Neverthenonetheless, there will be cases where we *do* get conjugacy, as in the following example.

16.1.2 Example 2

Consider the system

$$\begin{aligned} x' &= x + y^2 \\ y' &= -y \end{aligned}$$

Then $f(x,y) = (x+y^2, -y)$. Clearly we have only one equilibrium point, which is at the origin again. Thus the linearisation around it is given by

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

Without any further work, we can conclude that the linearised system is unstable and has a saddle. Note that if y is small, then y^2 would be even smaller. Thus we would expect the non-linear system to look quite similar to its linearisation around the origin.

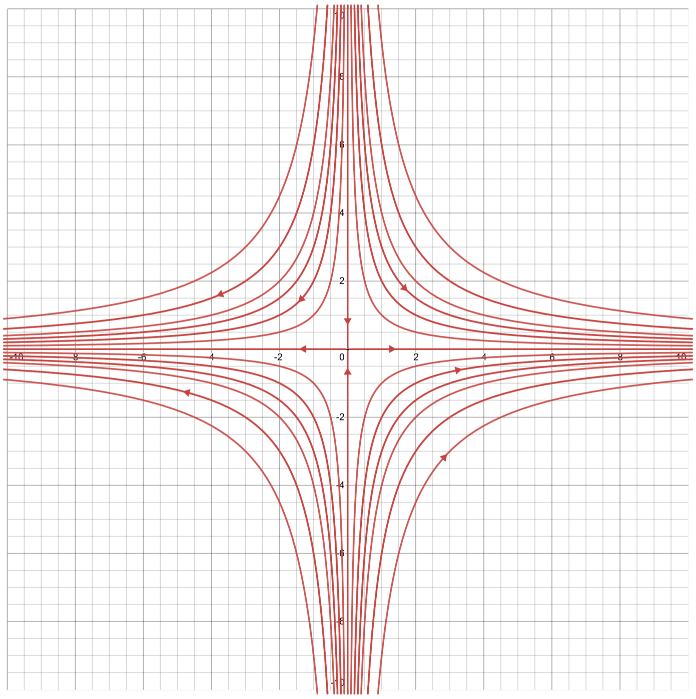

Figure 12: Phase portrait of linearised system, source

In this case we can actually solve the nonlinear system, allowing us to test our hypotheses. We can easily solve for y to get

$$y(t) = y_0 e^{-t}$$

Then we need to solve

$$x' = x + y_0^2 e^{-2t}$$

We know that the general solution is going to be of the form

$$x(t) = ce^t + x_p(t)$$

where $x_p(t)$ is a particular solution. We will guess x_p to be of the same form as the inhomogeneity (this is often not a bad first guess), so in particular we will guess

$$x_p(t) = be^{-2t}$$

where b is a constant to be determined. Substituting this into the ODE (for x), we find

$$b = -\frac{y_0^2}{3}$$

thus

$$x(t) = \left(x_0 + \frac{y_0^2}{3}\right)e^t - \frac{y_0^2}{3}e^{-2t}$$

(the coefficient of e^t is of that form so that $x(0) = x_0$). Plotting this we get the phase portrait in Figure 13.

The parabola going through the origin is given by $x = -\frac{1}{3}y^2$. If our initial conditions lie on this parabola, we get

$$x(t) = -\frac{y_0^2}{3}e^{-2t}, y(t) = y_0 e^t$$

which draws out the above parabola. Because solutions starting on this curve always remain on it, we call it a *stable curve*.

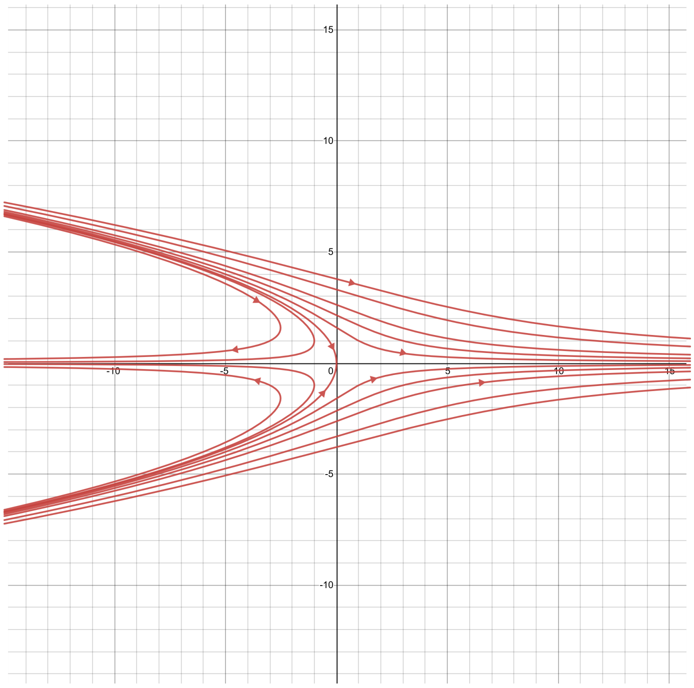

Figure 13: Phase portrait of non-linear system, source

Comparing phase portraits of the two systems, we can maybe convince ourselves of a certain similarity. The hope is that there is a way of translating one to the other (this is after all exactly what the homeomorphism would do). In fact in this case we can check that the change of variables given by

$$T(x, y) = \left(x + \frac{1}{3}y^2, y\right)$$

maps the flow of the linear system onto the non-linear one.

$$\begin{aligned}
T^{-1}e^{tA}T(x_0, y_0) &= T^{-1}e^{tA}\left(x_0 + \frac{1}{3}y_0^2, y_0\right) \\
&= T^{-1}\left(e^t\left(x_0 + \frac{1}{3}y_0^2\right), y_0e^{-t}\right) \\
&= \left(e^t\left(x_0 + \frac{1}{3}y_0^2\right) - \frac{y_0^2}{3}e^{-2t}, y_0e^{-t}\right)
\end{aligned}$$

which we know is exactly the flow of the non-linear system. In the case the map T was linearised the system globally and was a diffeomorphism. Most of the time, this is not the case; usually, we only get local homeomorphisms.