M2AA1 - ODEs - Revision Notes

Harry Gulliver

May 16, 2013

Contents

1 Existence and Uniqueness for Initial Value Problems

DEFINITION 1.1 (Norm)

Given a vector space V, a norm on V is a function $||.|| : V \to [0,\infty)$ satisfying the properties of:
$||\vec{v}|| > 0 \; \forall \vec{v} \neq \vec{0}$ *and* $||\vec{0}|| = 0$ *(positive-definiteness)*
$||c\vec{v}|| = |c| \times ||\vec{v}||$
$||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$ *(triangle inequality).*

A common norm used on function spaces is the supremum norm. If f is a function from an interval I to $\mathbb{R}$, then:

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

DEFINITION 1.2 (Banach Space)

A Banach space is a vector space with a norm that is complete. That is to say, any Cauchy sequence $\{\vec{x}_n\}$ converges to some $\vec{x}$ in the space.

The space of continuous functions from a compact (closed and bounded) interval to $\mathbb{R}$ is a Banach space wrt the supremum norm.

THEOREM 1.1 (Banach Contraction Theorem)

Let X be a Banach space and consider a function $f : X \to X$ with the property:

$$\exists \lambda \in (0,1) \; \forall x, y \in X \; ||f(x) - f(y)|| \leq \lambda ||x - y||$$

(such a function is called a contraction - for it brings points closer together).

Then f has a unique fixed point. That is:

$$\exists! p \in X \; f(p) = p$$

PROOF:

Pick any $x_0 \in X$ and define a sequence by $x_{n+1} = f(x_n)$. As f is a contraction, this sequence is Cauchy:

$$\begin{aligned} ||x_{n+1} - x_n|| &= ||f(x_n) - f(x_{n+1})|| \\ &\leq \lambda ||x_n - x_{n-1}|| \\ &\dots \\ &\leq \lambda^n ||x_1 - x_0|| \\ \Rightarrow \quad \forall n \geq m \; ||x_n - x_m|| &= ||x_n - x_{n-1} + x_{n-1} - x_{n-2} + \ldots + x_{m+1} - x_m|| \\ &\leq ||x_n - x_{n-1}|| + ||x_{n-1} - x_{n-2}|| + \ldots + ||x_{m+1} - x_m|| \\ &\leq (\lambda^{n-1} + \ldots + \lambda^m)||x_1 - x_0|| \\ &\leq \frac{\lambda^m}{1-\lambda}||x_1 - x_0|| \end{aligned}$$

So the sequence has a limit (it is a Cauchy sequence in a Banach space). Call this limit p:

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n \geq N \;\; ||x_n - p|| < \varepsilon$$

But, by the definition of x_n and the fact that f is a contraction:

$$||f(x_n) - f(p)|| \leq \lambda ||x_n - p|| < ||x_n - p|| < \varepsilon$$

So we see that $f(x_n) \to f(p)$. But $f(x_n) = x_{n+1}$ and $\lim_{n\to\infty}(x_n) = \lim_{n\to\infty}(x_{n+1})$, so $f(x_n)$ and x_n have the same limit. I.e., $f(p) = p$. So p is a fixed point.

We now prove uniqueness. Suppose $p \neq q$ are both fixed points of f. Then:

$$\begin{aligned} ||p - q|| &= ||f(p) - f(q)|| \\ &\leq \lambda ||p - q|| \\ \Rightarrow \quad (1-\lambda)||p - q|| &\leq 0 \end{aligned}$$

But $\lambda \in (0,1)$ and $||p-q|| \geq 0$, so $(1-\lambda)||p-q|| \geq 0$. Therefore, combining these two inequalities and dividing by the (non-zero) factor of $(1-\lambda)$, we have: $0 \leq ||p-q|| \leq 0$. Hence $||p-q|| = 0$. But, by the positive-definiteness of the norm, this is only possible if $p = q$. Contradiction.

■

DEFINITION 1.3 (Lipschitz Functions)

A function f is called Lipschitz if:

$$\exists K > 0 \;\; \forall x, y \;\; |f(x) - f(y)| \leq K|x - y|$$

The constant K is called the Lipschitz constant for the function. Functions may be locally Lipschitz (if the above condition is satisfied in some simply-connected, open subset of the domain) or globally Lipschitz (if the above condition is satisfied everywhere).

Note that a contraction on a Banach space, as defined above, is simply a Lipschitz function with $K < 1$.

Note also that any function differentiable on an interval (a,b) is automatically Lipschitz (at least locally) by the Mean Value Theorem. For take any $x, y \in (a,b)$; then:

$$\exists c \in (x,y) \;\; f'(c)[x-y] = f(x) - f(y) \quad \Rightarrow \quad |f(x) - f(y)| = |f'(c)| \times |x - y|$$

So we need simply take our Lipschitz constant K to be the supremum norm of f' on (a,b). This also holds for functions from $\mathbb{R}^n \to \mathbb{R}$, as, writing $J(x,t)$ as the Jacobian matrix of f:

$$|f(x,t) - f(y,t)| \leq \left|\left| J(\xi, t) \left(\frac{y-x}{||y-x||} \right) \right|\right| ||x - y||$$

for some ξ on the line segment joining x and y.

THEOREM 1.2 (Picard's Existence & Uniqueness Theorem)

Let U be an open subset of $\mathbb{R}^n \times \mathbb{R}$ containing a point $(x_0, 0)$ and consider a continuous function $f : U \to \mathbb{R} : (x,t) \mapsto f(x,t)$ which is Lipschitz in x. That is, for any fixed t, it satisfies the Lipschitz condition for its first argument; so there exists some K so that for all $x, y \in \mathbb{R}^n$: $|f(x,t) - f(y,t)| \leq K||x-y||$. Suppose also that f is bounded: $\exists M > 0 \ |f(x,t)| \leq M$.

Take $h \in \left(0, \frac{1}{2K}\right)$ sufficiently small that $[-h,+h] \times \{y \in \mathbb{R}^n : ||x_0 - y|| \leq hM\} \subseteq U$ (possible by the openness of U). Then there is a unique function $x : (-h,+h) \to \mathbb{R}^n : t \mapsto x(t)$ satisfying the initial value problem:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t) \qquad x(0) = x_0$$

PROOF:

Firstly, we rewrite our initial value problem as an integral equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t) \quad \text{and} \quad x(0) = x_0 \quad \Leftrightarrow \quad x(t) = x_0 + \int_0^t f(x(s),s)\,\mathrm{d}s$$

Let C be the space of continuous functions $x : (-h,+h) \to \mathbb{R}^n : t \mapsto x(t)$ (with h defined as in the statement of the theorem). Define a map $P : C \to C$ by:

$$P[x(t)] = x_0 + \int_0^t f(x(s),s)\,\mathrm{d}s$$

Then the solution of our initial value problem is equivalent to finding a fixed point of P: $x = P(x)$.

Note that:

$$\begin{aligned}
|P(x) - x_0| &= \left| \int_0^t f(x(s),s)\,\mathrm{d}s \right| \\
&\leq \int_0^t |f(x(s),s)|\,\mathrm{d}s \\
&\leq \int_0^h |f(x(s),s)|\,\mathrm{d}s \\
&\leq h \sup_{s \in (-h,+h)} \{|f(x(s),s)|\} \\
&\leq hM
\end{aligned}$$

So if $\forall t \in (-h,+h)\ x(t) \in B = \{y \in \mathbb{R}^n : ||x_0 - y|| \leq hM\}$, we see that $\forall t \in (-h,+h)\ P[x(t)] \in B \subseteq U$. This is necessary because f is only defined on U. So this proves that P is well-defined.

We now prove that P is a contraction on B. Take functions $y, z \in B\ \forall t \in (-h,+h)$. Then for any $t \in (-h,+h)$:

$$\begin{aligned}
|P[y(t)] - P[z(t)]| &= \left| \int_0^t [f(y(s),s) - f(z(s),s)]\ \mathrm{d}s \right| \\
&\leq \int_0^t |f(y(s),s) - f(z(s),s)|\ \mathrm{d}s \\
&\leq K \int_0^t |y(s) - z(s)|\,\mathrm{d}s \quad \text{Lipschitz} \\
&\leq K \int_0^h |y(s) - z(s)|\,\mathrm{d}s \quad h > t \text{ and integrand positive} \\
&\leq hK||y - z||_\infty \\
&\leq \frac{1}{2}||y - z||_\infty \quad \text{definition of } h < \tfrac{1}{2K}
\end{aligned}$$

But this holds $\forall t \in (-h,+h)$, so it holds also for the supremum norm of the lhs:

$$||P[y(t)] - P[z(t)]||_\infty \leq \frac{1}{2}||y - z||_\infty$$

So P is a contraction, as claimed. Hence, by the Banach contraction theorem, there is a unique fixed point of P; that is, a unique $x : (-h,+h) \to \mathbb{R} : t \mapsto x(t)$ such that:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t) \qquad x(0) = x_0$$

■

Some Remarks

There is a lot of detail in the statement of Picard's theorem, which can be confusing. At its simplest, the theorem states that if $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is Lipschitz in x and bounded, then $\mathrm{d}x/\mathrm{d}t = f(x,t)$, $x(0) = x_0$ has a unique solution. The definition of U is to allow the more general case where f is not defined on all of $\mathbb{R}^n \times \mathbb{R}$ but only a subset. The construction of the set B and attendant requirement that h be sufficiently small is to ensure that the problem remains well-defined in this case.

Picard's theorem is stated in terms of first-order initial value problems. However, it also provides proof of uniqueness for higher order, linear equations, by virtue of the fact that it applies to first-order systems as well as first-order, single-variable equations (as f is defined for x in $\mathbb{R}^n$). For consider the equation:

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_0(t)y = b(t)$$

This can be rewritten as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \\ y^{(n)} \end{pmatrix} = \begin{pmatrix} y' \\ y'' \\ \vdots \\ y^{(n)} \\ b(t) - \left[a_{n-1}(t)y^{(n-1)} + \ldots + a_0(t)y\right] \end{pmatrix}$$

Which, by Picard's theorem, has a unique solution.

The theorem is stated here in terms of an initial value problem at $t = 0$. Of course, there is nothing special about this starting point; it was chosen merely for ease of exposition. More generally, the theorem applies when the initial condition of the ODE is $x(\alpha) = x_0$ for any $(\alpha, x_0) \in U$.

The theorem holds for f locally Lipschitz, but is somewhat stronger if f is globally Lipschitz. The reason for this is that h depends on the Lipschitz constant, which is the same over the entire domain for a globally Lipschitz function, but varies for a locally Lipschitz function. Picard's theorem only guarantees existence and uniqueness of solutions in a small interval $(\alpha - h, \alpha + h)$ around the initial point. It can, however, be easily extended to cover a wider range. First consider the case of f globally Lipschitz. Suppose x_0 is our unique solution in $(-h,+h)$ of $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t)$, $x(0) = a$. Take $h_n = \frac{hn}{2}$ and consider the initial value problem $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t)$ and $x(h_1) = x_0(h_1)$. By Picard's theorem, there is a unique solution $x_1 : (h_1 - h, h_1 + h) \to \mathbb{R}$. But the domains of uniqueness of x_0 and x_1 intersect in $(-h/2,+h)$; so in this interval, by uniqueness, $x_0(t) = x_1(t)$.

Continuing in this way, we define x_n to be the unique solution in $(h_n - h, h_n + h)$ to $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t)$, $x(h_n) = x_{n-1}(h_n)$. Then x_n and x_{n+1} have overlapping domains of existence (both are defined on $(h_{n+1} - h, h_n + h)$ and so are equal in this interval, by uniqueness. So we can "piece together" a solution $x(t) = x_n(t)$ when $t \in (h_n - h, h_n + h)$. Allowing n to take negative integer values as well as positive, we can cover all of $\mathbb{R}$ in intervals of this form, in which x must be the unique solution to our initial value problem. So the problem has a unique solution for all time.

Now consider the case of f locally Lipschitz. The above argument still holds, in essence, but with the caveät that h is no longer a single constant value. It depends on the Lipschitz constant, which varies depending on which locality of $\mathbb{R}$ we are in. Because of this, h can "shrink" as we move along, from interval to interval. This means that we might not be able to cover the whole real line in our intervals. As such, existence and uniqueness of a solution to an initial value problem with a locally Lipschitz function is

generally only guaranteed for finite time. The above argument, while generally incapable of covering the whole real line, can at least find us the maximal interval on which the solution has guaranteed existence and uniqueness. This maximal interval can, however, be more easily found by examining when the solution tends to $\pm\infty$, which must always happen in finite time for such solutions.

Finally, a sometimes useful consequence of the method of proof of Picard's theorem is that it gives us a method for constructing the unique solution, called Picard iteration. To solve the problem $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t)$, $x(0) = \chi$, take $x_0(t) \equiv \chi$ and inductively define:

$$x_{n+1}(t) = \chi + \int_0^t f(x_n(s), s)\,\mathrm{d}s$$

Then, by the details of the proof of the Picard and Banach contraction theorems, the sequence $\{x_n\}$ converges to the unique solution. This might not always be the easiest way to find a solution, but in principle it is guaranteed to work.

THEOREM 1.3 (Continuous Dependence on Initial Conditions and the Butterfly Effect)

Let $U \subseteq \mathbb{R}^n \times \mathbb{R}$ be open, $f, g : U \to \mathbb{R}^n$ continuous, such that:

$$K = \left|\left| \frac{f(x,t) - f(y,t)}{x - y} \right|\right|_\infty, \quad M = ||f(x,t) - g(x,t)||_\infty$$

are both finite. Let x, y be respective solutions of the initial value problems:

$$\begin{cases} \frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t) \\ x(0) = x_0 \end{cases} \qquad \begin{cases} \frac{\mathrm{d}y}{\mathrm{d}t} = g(y,t) \\ y(0) = y_0 \end{cases}$$

Then:

$$|x(t) - y(t)| \leq |x_0 - y_0| e^{K|t|} + \frac{M}{K}(e^{K|t|} - 1)$$

In particular, for $f \equiv g$, $M = 0$, so we have:

$$|x(t) - y(t)| \leq |x_0 - y_0| e^{K|t|}$$

This shows that the distance between solutions for two different initial values of the same ODE depends on the distance between their initial values; so after a fixed time, the closer the initial values, the closer we should expect the solutions to be. However, the solutions may separate up to exponentially quickly, so that, even after a very short time, the solutions may be very far apart. This is the well-known "butterfly effect".

THEOREM 1.4 (Flow)

When the function f in an IVP does not depend explicitly on t, it can often be convenient to represent the solution to the IVP $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x)$ $x(0) = x_0$ as $\phi_t(x_0)$. That is, $\frac{\mathrm{d}\phi_t(x)}{\mathrm{d}t} = f(\phi_t(x))$ and $\phi_0(x) = x$. Such IVPs are called autonomous. This highlights the important flow property for any IVP with a unique solution:

$$\phi_{s+t}(x) = \phi_t[\phi_s(x)]$$

This is in some sense intuitively obvious; if you start at x and follow a path for time $s+t$, you will arrive at the same point as if you start at x, follow the path for time s, then, starting from there, follow the path for time t.

PROOF:

Let $\psi : a \mapsto \phi_{s+a}(x)$. So $\phi_{s+t}(x) = \psi(t)$ and $\phi_s(x) = \psi(0)$. So we wish to show that $\psi(t) = \phi_t[\psi(0)]$

Consider the IVP $\frac{\mathrm{d}y}{\mathrm{d}t} = f(y)$, $y(0) = \psi(0)$. This has a unique solution $\phi_t[\psi(0)]$. But $\psi(t)$ is also a solution, for it clearly satisfies the initial condition and $\frac{\mathrm{d}\psi(t)}{\mathrm{d}t} = \frac{\mathrm{d}\phi_{s+t}(x)}{\mathrm{d}t} = f[\phi_{s+t}(x)] = f[\psi(t)]$. So $\psi(t)$ and $\phi_t[\psi(0)]$ are both solutions. Hence, by uniqueness, they are the same:

$$\phi_{s+t}(x) = \phi_t[\phi_s(x)]$$

■

2 Linear Systems

DEFINITION 2.1 (Matrix Exponentiation)

Let $M \in \mathbb{R}^{n\times n}$. Define $M^0 = I_n$, the $n \times n$ identity matrix. Then, by analogy with the power series definition of exponentiation in $\mathbb{C}$, define the exponential of M to be:

$$e^M = \sum_{r=0}^{\infty} \frac{1}{r!} M^r$$

Similarly, define the exponential of Mt, for some scalar t, as:

$$e^{Mt} = \sum_{r=0}^{\infty} \frac{1}{r!} (tM)^r$$

The powers of M and sums within the power series are all well-defined, as M is $n \times n$, so it only remains to determine whether or not the sum converges. This is shown by the comparison test, defining the norm of a matrix M to be the maximum of the absolute values of its coefficients and comparing with $e^{n||M||}$, which converges, hence $\exp(M)$ converges absolutely. Note that e^M is a matrix of the same dimensions as M.

PROPOSITION 2.1 (Useful Results)

Let A, B, T be $n \times n$ matrices, with T invertible. Then:

1.
$$\exp(T^{-1}AT) = T^{-1}\exp(A)T$$

2.
$$BA = AB \quad \Rightarrow \quad \exp(A+B) = \exp(A)\exp(B)$$

3.
$$\exp(-A) = [\exp(A)]^{-1}$$

PROPOSITION 2.2 (Derivative of Matrix Exponentials)

$$\frac{\mathrm{d}e^{Mt}}{\mathrm{d}t} = Me^{Mt} = e^{Mt}M$$

PROOF:

$$\begin{aligned}
\frac{e^{M(t+h)} - e^{Mt}}{h} &= \frac{e^{Mt}e^{Mh} - e^{Mt}}{h} \\
&= e^{Mt}\frac{e^{Mh} - I}{h} \\
&= e^{Mt}\frac{1}{h}\sum_{r=1}^{\infty}\frac{1}{r!}(Mh)^r \\
&= e^{Mt}\sum_{r=1}^{\infty}\frac{1}{r!}M^r h^{r-1} \\
&\to e^{Mt}M
\end{aligned}$$

And the proof of commutativity follows from part 2 of the above proposition.

■

It is then easy to check from this that $e^{Mt}\vec{x}_0$ is a solution to the IVP $\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = M\vec{x}$, $\vec{x}(0) = \vec{x}_0$. By Picard's theorem, this is the unique solution. So any general linear system can be solved in terms of matrix exponentials.

The problem with this is that it is not necessarily easy to evaluate the exponential. The power series definition is rather unwieldy and difficult to compute. The Jordan Canonical Form Theorem from Linear Algebra is useful for this.

DEFINITION 2.2 (Eigenvalues and Eigenvectors)

Recall that a vector v is an eigenvector of the square matrix A iff $v \neq 0$ and $\exists\lambda \;\; Av = \lambda v$. This λ is called the eigenvalue of A corresponding to v. The eigenvalues are found as the roots of the characteristic polynomial, $det(A - \lambda I)$

The multiplicity with which an eigenvalue occurs as a root of the characteristic polynomial is its algebraic multiplicity. The dimension of the eigenspace (the space of vectors w mapped to λw by A) is called the geometric multiplicity.

If there are n distinct eigenvalues, then there are n linearly independent eigenvectors and, letting P be the matrix whose columns are the eigenvectors, P is invertible and $P^{-1}AP$ is diagonal, with the eigenvalues as diagonal entries. Raising a diagonal matrix to a power is easy - you just raise the individual entries to the power - so it is easy to prove the the exponential of a diagonal matrix D is the diagonal matrix whose diagonal entries are the exponentials of the entries of D. Then, by part 1 of the above proposition, the matrix is easy to exponentiate.

If the matrix to be exponentiated does not have a full set of linearly independent eigenvectors, then the Jordan canonical form can be used to simplify the calculation of the exponential. Recall:

DEFINITION 2.3 (Direct Sums)

The direct sum of two matrices, $A \oplus B$, is the matrix with A and B arranged on the leading diagonal and zeroes everywhere else. E.g.:

$$\begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} \oplus \begin{pmatrix} 7 & 2 \\ 1 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 7 & 2 \\ 0 & 0 & 1 & 8 \end{pmatrix}$$

DEFINITION 2.4 (Jordan Blocks)

The Jordan block $J_n(\lambda)$ is $n \times n$ matrix with λ on all leading diagonal entries, 1 on the superdiagonal and 0 everywhere else. E.g.:

$$J_3(4) = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}$$

DEFINITION 2.5 (Similarity) *Two matrices, A and B are said to be similar if there exists an invertible matrix P such that $P^{-1}AP = B$. Similarity is an equivalence relation.*

THEOREM 2.1 (Jordan Canonical Form Theorem)

Any matrix A with complex coefficients (of course, real coefficients can be regarded as complex with zero imaginary part) is similar to a matrix in Jordan form, which is a direct sum of Jordan blocks $J_{a_i}(\lambda_i)$ where λ_i is the i^{th} eigenvalue.

A matrix expressed in Jordan form as J can generally be exponentiated as follows: write it as $J = L + N$ where L is the diagonal matrix consisting of the eigenvalues and N is the upper triangular matrix consisting of 1's on the superdiagonal and zeroes elsewhere. Check that $LN = NL$. Then note that $N^m = 0$ for some sufficiently large m, so that the power series of e^N is finite and thus easily calculable. L is diagonal and so easy to exponentiate. So exponentiate both L and N separately, then use $e^{L+N} = e^L e^N$. We can then reverse the similarity transformation to obtain the exponential of the original matrix.

3 Power-Series Solutions

Suppose we wish to solve a differential equation $y' = f(x,y)$. Supposing y to be analytic (this is justifiable so long as f is analytic - proof beyond the scope of this course), we can write y as a power series, $y = \sum_{n=0}^{\infty} a_n x^n$. Substituting this into the ODE will give us an equation that can then (in many cases) be solved for the coefficients a_n by comparing powers of x, giving us a solution in the form of a power series (which may be recognisable as a known function).

DEFINITION 3.1 (Poles)

A pole of a function is a point at which that function is not defined; typically this arises because of division by zero. For instance, $1/x$ has a pole at $x = 0$. If a function $f(x)$ has a pole at $x = x_0$, but there is some integer m such that $(x-x_0)^m f(x)$ does not have a pole there, then the pole is said to have order m (where m is the least such integer). For instance, if $f(x) = 1/x$, $xf(x) = 1$, which does not have a pole at $x = 0$, so this pole of $f(x)$ has order 1.

If a second-order, linear ODE can be written as $y'' + py' + qy = 0$, where p has a pole of order at most 1 and q has a pole of order at most 2, then the points at which these poles occur are called regular poles or regular singular points.

The Picard theorem does not pertain to equations with regular singular points, but it can be shown (beyond the scope of this course) that a solution to such a regular equation always does exist and has the form $y = x^\alpha \sum_{n=0}^{\infty} a_n x^n$ where $\alpha \in \mathbb{R}\backslash\mathbb{Z}$.

As with a normal power series solution, we can substitute this into the ODE, compare coefficients of terms of equal order, and arrive at equations for the coefficients a_n. In this case, we will also obtain an equation for α from the a_0 term (w.l.o.g., we assume this is non-zero; otherwise, let m be the index of the first non-zero term and let $\alpha' = \alpha + m$, then proceed with this series). In fact, we will generally find that this equation is a quadratic for second-order equations such as we consider. If it has two distinct roots and these do not differ by an integer, then this gives us two distinct, linearly independent solutions. Otherwise, if we have one solution, y and it can be shown that $log(x)y$ is the second solution.

4 Boundary Value Problems

We have a powerful result concerning the existence and uniqueness of solutions to initial value problems. Let us now consider boundary value problems, which can be rather more subtle. Consider the ODE $y'' + y = 0$, which has general solution $y = A\cos(x) + B\sin(x)$. Take the boundary conditions $y(0) = 0$, $y(\pi) = 0$. We find $A = 0$, but cannot solve for B. $y = B\sin(x)$ satisfies this BVP for *any* real B. Consider instead the same equation with boundary conditions $y(0) = 0, y(\pi) = a \neq 0$. We again find that $A = 0$, but now cannot find any value of B which will solve the BVP. So a perfectly mundane looking ODE, which would, by Picard, have had guaranteed unique solutions if given any initial conditions $y(0) = y_0, y'(0) = y_0'$, can have uncountably many or no solutions for certain choices of boundary conditions. Boundary conditions could also be chosen which would give a unique solution, such as $y(0) = 0$, $y(\pi/2) = 1$, which gives $A = 0$, $B = 1$.

Consider the equation $y'' - \lambda y = 0$, where λ is a free parameter, with boundary conditions $y(0) = y(\pi) = 0$. We will proceed with the analysis by considering 3 cases:

1: $\lambda = 0$:

$y = Ax + B$. This can only satisfy our boundary conditions for $A = B = 0$, so we only have the trivial solution $y \equiv 0$.

2: $\lambda = \mu^2 > 0$:

$y = A\cosh(\mu x) + B\sinh(\mu x)$. $y(0) = 0$ gives $A = 0$, but then $y(\pi) = 0$ gives $B = 0$, so again we only get the trivial solution.

3: $\lambda = -\mu^2 < 0$:

$y = A\cos(\mu x) + B\sin(\mu x)$. $y(0) = 0$ gives $A = 0$, then $y(\pi) = 0$ gives us that $B\sin(\mu\pi) = 0$. This implies either $B = 0$, leaving us with the trivial solution yet again, or $\mu \in \mathbb{Z}$, with B able to take any value.

So, if we neglect the trivial solution, we can only obtain solutions for $\lambda = -n^2$ and these solutions are $y = B_n \sin(nx)$ where, w.l.o.g., we take $n \in \mathbb{N}$ (using parity of $\sin(x)$). Because this is a linear, homogenous ODE, solutions sum, so the general solution is:

$$y = \sum_{n=1}^{\infty} B_n \sin(nx)$$

So long as this sum converges and is twice differentiable.

Extra conditions will be needed to determine the coefficients B_n. In practice, this sort of solution typically arises in the study of PDEs such as the heat and wave equations, where an initial condition will furnish the coefficients through Fourier analysis.

Consider the equation $y'' = \lambda y$ analysed above. In terms of the linear, differential operator $\mathscr{L} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}$, this can be written $\mathscr{L}(y) = \lambda y$. So the ODE can be regarded as finding an eigenvector of a linear operator in an abstract function space. In this context, we use the term "eigenfunction" rather than "eigenvector". So, with the boundary conditions we had above, $\lambda_n = -n^2$ is an eigenvalue with eigenfunction $y_n = \sin(nx)$.

The above example is a special case of a type of equation called a Sturm-Liouville problem. The general form of these problems is:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right) + q(x)y + \lambda r(x)y = 0$$

with boundary conditions $\alpha_0 y(a) + \alpha_1 y'(a) = 0$, $\beta_0 y(b) + \beta_1 y'(b) = 0$. λ is a free parameter.

THEOREM 4.1 (Sturm-Liouville Theorem)

Let $p(x) > 0$, $r(x) > 0$ be continuous and p continuously differentiable on $[a,b]$. Then the Sturm-Liouville problem:

$$(py')' + qy + \lambda ry = 0$$

with boundary conditions as above has infinitely many eigenvalues λ, which are all real, all have single multiplicity and which, if ordered so that $\lambda_n < \lambda_{n+1}$, are such that $\lambda_n \to \infty$.

Furthermore, the corresponding eigenfunctions are orthogonal under an integral inner product with $r(x)$ as weight function:

$$\forall m \neq n \quad \int_a^b y_n(x) y_m(x) r(x)\,\mathrm{d}x = 0$$

Finally, every C^2 (continuously twice-differentiable) function on $[a,b]$ can be written as a series in the eigenfunctions:

$$f(x) = \sum_{n=1}^{\infty} A_n y_n(x)$$

where:

$$A_n = \frac{\int_a^b f(x) y_n(x) r(x)\,\mathrm{d}x}{\int_a^b [y_n(x)]^2 r(x)\,\mathrm{d}x}$$

This last, key point of the theorem should be compared with Fourier expansions. In the special case $p \equiv 1$, $q \equiv 0$, $r \equiv 1$, the eigenfunctions are the Fourier basis functions, $\sin(nx)$ and $\cos(nx)$, and their norm (the weighted integral of the square of the function, in the denominator) is $(b-a)/2$, so that the above ratio of integrals reduces to our familiar formula for Fourier coëfficients. Thus, the Sturm-Liouville theorem lets us perform Fourier-like expansions of functions in terms of other sequences of orthogonal functions, found as the roots of Sturm-Liouville problems.

DEFINITION 4.1 (The Wronskian)

Let y, z be solutions of a second-order ODE. Then the Wronskian W is defined as follows:

$$\begin{aligned} W(x) &= \det\begin{pmatrix} y & y' \\ z & z' \end{pmatrix} \\ &= y(x)z'(x) - y'(x)z(x) \end{aligned}$$

PROPOSITION 4.1 (Sign of the Wronskian)

When considering an ODE of the form $(py')' + ry = 0$ where $p > 0$ is continuously differentiable, the Wronskian has constant sign and is zero iff y, z are linearly dependent.

Of course, the determinant of a matrix is negated if the rows are swapped, so the particular sign depends on the order in which we take y and z.

PROOF:

We have: $(py')' + ry = 0$ and $(pz')' + rz = 0$. Multiplying the first equation by z and the second by y and subtracting, we get:

$$pyz'' - pzy'' + p'z'y - p'y'z + ryz - ryz = 0$$

Which simplifies to:

$$p(zy'' - yz'') + p'W = 0$$

Note that:

$$\begin{aligned} W'(x) &= (yz'' + y'z') - (y'z' + y''z)zz \\ &= yz'' - y''z \end{aligned}$$

Hence, substituting in the above, we see:

$$pW' + p'W = 0$$

I.e., $(pW)' = 0$. But this holds for all x, so we conclude that pW is constant. p is assumed strictly positive, so $W = c/p$ has constant sign.

The fact that $W = 0$ iff y and z are linearly dependent follows immediately from the definition of the Wronskian as the determinant of a matrix.

■

THEOREM 4.2 (Sturm Separation Theorem)

Let y, z be two linearly independent solutions to $(py')' + ry = 0$. Then the zeroes of the two solutions alternate. That is, given a, b such that $y(a) = y(b) = 0$ and $y(x) \neq 0 \ \forall x \in (a,b)$, we have $\exists! c \in (a,b) \ z(c) = 0$ and vice versa.

PROOF:

Take $a < b$ consecutive zeroes of y. Then $W(a) = y(a)z'(a) - y'(a)z(a) = -y'(a)z(a)$ and $W(b) = -y'(b)z(b)$. As the Wronskian has constant sign, $y'(a)z(a)$ and $y'(b)z(b)$ must have the same sign. But $y'(a)$ and $y'(b)$ must have opposite sign (since they are consecutive zeroes), so $z(a)$ and $z(b)$ must have opposite sign too. Hence, by the intermediate value theorem, there is some $c \in (a,b)$ such that $z(c) = 0$. Furthermore, c must be unique; for if not, there are two zeroes, c and d, of z between a and b, so by identical reasoning, y has a zero in (c,d); but this contradicts the assumption that a and b are *consecutive* zeroes of y.

■

5 Some Useful Multivariable Calculus Results

DEFINITION 5.1 (The Jacobian of a Function)

Let $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$, $f : U \to V$. We say f is differentiable at $x \in U$ if there is a matrix $J \in \mathbb{R}^{m \times n}$ such that:

$$\lim_{||u|| \to 0} \left[\frac{||f(x+u) - f(x) - Ju||}{||u||} \right] = 0$$

The matrix J is called the Jacobian of f at x and is often written $J = Df_x$ or $J = Df(x)$.

It can easily be shown that the components of Df_x are the partial derivatives: $Df_x = (a_{i,j})$ where: $a_{i,j} = \frac{\partial f(x)_i}{\partial x_j}$

A generalisation of the Taylor expansion theorem allows us to write:

$$f(x+u) = f(x) + Df_x u + O(||u||)$$

THEOREM 5.1 (The Inverse Function Theorem)

Let $U \subseteq \mathbb{R}^n$ be open, $p \in U$ and $f : U \to \mathbb{R}^n$ continuously differentiable. If Df_p is invertible, then there exist open sets $W \subseteq U$ and $V \subseteq \mathbb{R}^n$ with $p \in W$ and $f(p) \in V$ so that $f : W \to V$ is bijective and therefore has an inverse, $g : V \to W$, and g is differentiable, with $Dg_{f(p)} = (Df_p)^{-1}$.

In short, this theorem says that if a function from $\mathbb{R}^n$ to $\mathbb{R}^n$ has an invertible Jacobian at that point, then, locally to that point, the function is invertible and the Jacobian of the inverse at $f(p)$ is the inverse of the Jacobian of the original function at p. This is a generalisation of the theorem from single-variable real analysis that says that if $f'(p) \neq 0$, f has a local inverse g and g is differentiable at $f(p)$, with $g'(f(p)) = \frac{1}{f'(p)}$.

A differentiable function with a differentiable inverse is called a diffeomorphism.

THEOREM 5.2 (The Implicit Function Theorem)

Let $f : \mathbb{R}^p \times \mathbb{R}^n \to \mathbb{R}^n$ be differentiable, with $f(a,b) = 0$. Note that we consider $\mathbb{R}^p$ to be a parameter space, so that f is a function with p real parameters and n real arguments. Let J be the Jacobian of f with respect to its arguments, excluding its parameters. That is, take the Jacobian Df and remove the first p columns, which correspond to the parameters; the remaining $n \times n$ matrix is J. So J is the matrix of partial derivatives of f wrt its n 'proper' arguments, neglecting the parameters.

Then if J is invertible, there is a function $G : \mathbb{R}^p \to \mathbb{R}^n$ such that for all (x,y) in some neighbourhood of (a,b):

$$y = G(x) \quad \Leftrightarrow \quad f(x,y) = 0$$

That is, the zero-level curve of f, in the neighbourhood of (a,b), is a well-defined function of the parameters.

To illustrate the implicit function theorem, consider the case $f(x,y) = x^2 + y^2 - 1$. We have:

$$Df_x = (2x, 2y)$$

Therefore, if we treat x as a parameter and y as the 'true' variable, our derivative wrt the 'true' variable is $J = 2y$. Where this does not equal zero, i.e., for $y \neq 0$, we can solve the zero-level set of f for y as a function of x. The zero-level set is $x^2 + y^2 = 1$, the unit circle, so the theorem says that if we pick $y \neq 0$, there is a neighbourhood of y in which $x^2 + y^2 = 1$ iff $y = g(x)$. We can clearly see that, in this case, if we pick $y > 0$, we get $g(x) = \sqrt{1 - x^2}$ and the negative of this if $y < 0$. However, if $y = 0$, then $J = 0$ and we do not meet the conditions for the theorem; indeed, in a neighbourhood of $y = 0$, we cannot express y as a single-valued function of x along the unit circle.

6 Calculus of Variations

The subject of the calculus of variations concerns itself with a type of minimisation (or maximisation) problem. In elementary calculus, minimisation problems are concerned with finding a value of a variable which minimises a function. The calculus of variations is concerned with finding a *function* which minimises a "functional" - a function which acts on a function.

More specifically, define the functional:

$$I[y] = \int_0^1 f(x, y(x), y'(x)) \, \mathrm{d}x$$

where f is a known function and y a 'variable function', if you will - an undetermined function acted on by I. The functional, I, is here a map from a function space to the real line. The calculus of variations studies techniques for finding a function y which minimises I, subject to certain restrictions. Standard examples include finding the path between two points which minimises the travel time of an object moving along that path under gravity (the brachistochrone problem), finding the shape which maximises area for a given perimeter, &c.

THEOREM 6.1 (The Euler-Lagrange Equation)

Let $C = C^1([a,b], \mathbb{R})$ be the space of continuously differentiable functions $[a,b] \to \mathbb{R}$ and f some function $[a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Take a point $(x,y,z) \in [a,b] \times \mathbb{R}^n \times \mathbb{R}^n$ and denote by f_x, f_y, f_z the corresponding parts of the Jacobian matrix. That is, $f_x \in \mathbb{R}$, $f_y, f_z \in \mathbb{R}^n$ and the matrix $(f_x, f_y, f_z) \in \mathbb{R}^{2n+1}$ is the $(1 \times (1+n+n))$ Jacobian of f. Suppose f_y and f_z are continuous.

Define $I : C \to \mathbb{R}$ by:

$$I[y] = \int_a^b f[x, y(x), y'(x)] \, \mathrm{d}x$$

Given $Y \in C$, denote $f_y[Y](x) = f_y(x, Y(x), Y'(x))$ and $f_{y'}[Y](x) = f_z(x, Y(x), Y'(x))$. Fix $y_a, y_b \in \mathbb{R}^n$ and define:

$$A = \{y \in C \, : \, y(a) = y_a, \, y(b) = y_b\}$$

Then if $I : A \to \mathbb{R} : y \mapsto I[y]$ has a minimum at Y, then:

$$\forall v \in C : v(a) = v(b) = 0 \qquad \int_a^b \left[f_y[Y](x)v(x) + f_{y'}[Y](x)v'(x) \right] \mathrm{d}x = 0$$

and

$$\frac{\mathrm{d}f_{y'}[Y]}{\mathrm{d}x} = f_y[Y]$$

PROOF:

First note that, from the definitions of A and v, $\forall y \in A \ \ \forall h \in \mathbb{R} \ \ y + hv \in A$. Let Y be a minimiser of I in A and define $H : \mathbb{R} \to \mathbb{R} : h \mapsto I[Y + hv]$. Then H has a minimum at 0. So if H is differentiable, $H'(0) = 0$.

$$\begin{aligned}
\frac{H(h) - H(0)}{h} &= \int_a^b \frac{f[x, (Y+hv)(x), (Y+hv)'(x)] - f[x, Y(x), Y'(x)]}{h} \, \mathrm{d}x \\
&= \int_a^b \left(f_y[Y](x)v(x) + f_{y'}[Y](x)v'(x) + O(h) \right) \mathrm{d}x \quad \text{Taylor's theorem} \\
&\to \int_a^b \left(f_y[Y](x)v(x) + f_{y'}[Y](x)v'(x) \right) \mathrm{d}x \quad \text{as } h \to 0 \\
\int_a^b f_{y'}[Y](x)v'(x) \, \mathrm{d}x &= \left[f_{y'}[Y](x)v(x) \right]_a^b - \int_a^b \frac{\mathrm{d}f_{y'}[Y](x)}{\mathrm{d}x} v(x) \, \mathrm{d}x \\
&= -\int_a^b \frac{\mathrm{d}f_{y'}[Y](x)}{\mathrm{d}x} v(x) \, \mathrm{d}x \\
\Rightarrow \quad H'(0) &= \int_a^b \left(f_y[Y](x) - \frac{\mathrm{d}f_{y'}[Y](x)}{\mathrm{d}x} \right) v(x) \, \mathrm{d}x
\end{aligned}$$

So the derivative $H'(0)$ exists and has the above form. As Y is a minimiser, we have $H'(0) = 0$, so:

$$\int_a^b \left(f_y[Y](x) - \frac{\mathrm{d}f_{y'}[Y](x)}{\mathrm{d}x} \right) v(x) \, \mathrm{d}x = 0$$

This implies that the expression in parentheses is identically 0. For suppose otherwise, that there exists some $x_0 \in (a, b)$ such that:

$$G(x_0) = f_y[Y](x_0) - \frac{\mathrm{d}f_{y'}[Y]}{\mathrm{d}x}(x_0) \neq 0$$

W.l.o.g., suppose $G(x_0) > 0$. Then, by continuity, there is some neighbourhood of x_0 in which G is always positive. So choose v to be positive in this neighbourhood and zero everywhere else (recall that v is arbitrary apart from constraints on its endpoints). Then $\int_a^b G(x)v(x) \, \mathrm{d}x > 0$. Contradiction. So we conclude:

$$f_y[Y](x) = \frac{\mathrm{d}f_{y'}[Y](x)}{\mathrm{d}x}$$

■

Note that this is a necessary but not a sufficient condition. Just as a function from $\mathbb{R}^n$ to the real line may have a saddle point rather than an extremum where its derivative is zero, so too may a function satisfy the Euler-Lagrange equation without minimising or maximising the relevant functional. Furthermore, the existence of a minimiser/maximiser is by no means guaranteed. For the purposes of this course, however, such issues will be neglected.

In many cases, the function f has no explicit x-dependence. In this case, the Euler-Lagrange equation may usefully be simplified as follows:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} \left[f_{y'}[Y]Y' - f[Y(x)] \right] &= \frac{\mathrm{d}f_{y'}[Y](x)}{\mathrm{d}x} Y' + f_{y'}[Y]Y'' - \left[f_x[Y] + f_y[Y]Y' + f_{y'}[Y]Y'' \right] \\
&= Y' \left[\frac{\mathrm{d}f_{y'}[Y](x)}{\mathrm{d}x} - f_y[Y] \right] - f_x[Y]
\end{aligned}$$

By the Euler-Lagrange equation, the bracketed term is zero. As f has no explicit x-dependence, $f_x \equiv 0$. So we conclude:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f_{y'}[Y]Y' - f[Y(x)]\right] = 0$$

Hence:

$$f_{y'}[Y]Y' - f[Y] = C$$

This is often much simpler to solve than the full Euler-Lagrange equation.

THEOREM 6.2 (Lagrange Multiplier Theorem)

Let $g : \mathbb{R}^n \to \mathbb{R}$, $M = \{x \in \mathbb{R}^n : g(x) = 0\}$ be the zero-level set of g and $f : M \to \mathbb{R}$ have a minimum at X. Then if $Dg(X)$ has non-zero determinant, $\exists \lambda \in \mathbb{R}$ such that $Df(X) = \lambda Dg(X)$.

THEOREM 6.3 (Constrained Euler-Lagrange Equation)

Let $f, g : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and define functionals I, J on $A = \{y \in C^1([a,b],\mathbb{R}) : y(a) = y_a,\ y(b) = y_b\}$ as follows:

$$I[y(x)] = \int_a^b f[x, y(x), y'(x)]\ \mathrm{d}x$$

$$J[y(x)] = \int_a^b g[x, y(x), y'(x)]\ \mathrm{d}x$$

Suppose $Y \in A$ maximises (or minimises) I subject to $J[Y] = 0$. Then there exists some $\lambda \in \mathbb{R}$ such that the normal Euler-Lagrange condition holds for $F = f - \lambda g$. That is:

$$\frac{\mathrm{d}F_{y'}[Y]}{\mathrm{d}x} = F_y[Y]$$

7 Non-Linear Theory

In this section, we consider autonomous initial value problems $x' = f(x)$, $x(0) = x_0$ where $f : \mathbb{R}^n \to \mathbb{R}^n$ is infinitely differentiable. From section 1, we know that such an equation has a unique solution in some neighbourhood of 0; that there is a unique maximal somain of existence for the solution and a unique solution on this domain; that, if this domain is finite, $||x(t)|| \to \infty$ as t approaches the boundary of the domain; and the flow property: $\phi_{s+t}(x) = \phi_t[\phi_s(x)]$.

The curves $t \mapsto \phi_t(x)$ are called orbits. Rather than studying individual solutions for specific initial conditions, we will study the orbits generally, without specifying the value of the initial condition x.

DEFINITION 7.1 (Critical Points)

A point x such that $\phi_t(x) = x\ \forall t$ is called a critical point. These are also known as fixed points, singular points, singularities and rest points.

Even if an ODE is highly non-linear, it is often sufficient to consider a linear approximation to the equation near critical points. Let x_0 be a critical point. By Taylor's theorem, we write: $f(x) = f(x_0) + Df_{x_0}(x - x_0) + O(||x - x_0||^2)$. Suitably close to x_0, we can neglect high order terms in $||x - x_0||^2$ and use a linear approximation.

THEOREM 7.1 (Stable and Unstable Manifold Theorem)

If all the eigenvalues of $Df(x_0)$ have non-zero real part, $Df(x_0)$ is said to be hyperbolic. Order the eigenvalues $\lambda_1, \ldots, \lambda_n$ so that $Re(\lambda_1) < 0$ for $i \in \{1, \ldots, s\}$ and $Re(\lambda_i) > 0$ for $i \in \{s+1, \ldots, n\}$. Then there is a curve/surface/&c. $W^s(x_0)$, of dimension s, and a curve/surface/&c. $W^u(x_0)$, which intersect at x_0, such that:

$$x \in W^s(x_0) \quad \Rightarrow \quad \phi_t(x) \to x_0 \text{ as } t \to \infty$$

$$x \in W^u(x_0) \quad \Rightarrow \quad \phi_t(x) \to x_0 \text{ as } t \to -\infty$$

That is, there is some curve or similar along which solutions move towards x_0 and another along which they move away from x_0 and the dimensions of these curves/surfaces/whatever are the numbers of eigenvalues with negative and positive real parts respectively. W^s and W^u are called the stable and unstable manifolds, respectively.

If $s = n$, so that all orbits converge to x_0, then x_0 is called a sink.
If $1 \leq s < n$, so that some orbits converge, some flee, x_0 is called a saddle.
If $s = 0$, so that all orbits flee, x_0 is called a source.

THEOREM 7.2 (Hartman-Grobman)

Let x_0 be a singular point of the IVP $x' = f(x)$, with associated flow $\phi_t(x)$. Let $\psi_t(x)$ be the flow associated with $x' = Jx$, where $J = Df(x_0)$. If J is hyperbolic (all its eigenvalues have non-zero real part), then there is a continuous bijection $h : \mathbb{R}^n \to \mathbb{R}^n$ with a continuous inverse (a homeomorphism), such that, near x_0:

$$[\phi_t(x)] = h^{-1}(\psi_t[h(x)])$$

Think of this as taking the point x, mapping it to $h(x)$, then following the orbit of the linearised system from that point, then taking the inverse map back; this will give the same result as following the non-linearised orbit for the same amount of time.

This is useful because f might be a highly non-linear function, making the IVP $x' = f(x)$ difficult to analyse. $x' = Jx$, on the other hand, is a linear ODE (the linearisation of the original ODE), and so generally soluble analytically. This theorem tells us that there is a suitably nice function relating the solution of the non-linear ODE to the linearised ODE. In other words, a linear approximation to the ODE will give us accurate information about the behaviour of the non-linear ODE. In practice, we will not usually be able to find h, so we will not have exact information; however, as a result of the 'niceness' of h (it is a homeomorphism), the qualitative behaviour of the linear system will be the same as the non-linear system in some neighbourhood of the critical point.

DEFINITION 7.2 (Stability and Asymptotic Stability)

Let x_0 be a critical point. We call x_0 stable if:

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \ x \in B_\delta(x_0) \quad \Rightarrow \quad \phi_t(x) \in B_\varepsilon(x_0) \ \ \forall t \geq 0$$

That is, if we start close enough to the critical point, we never drift too far away.

x_0 is asymptotically stable if:

$$\exists \varepsilon > 0 \ \ x \in B_\varepsilon(x_0) \quad \Rightarrow \quad \phi_t(x) \to x_0 \text{ as } t \to \infty$$

That is, if all orbits that start close to x_0 converge to x_0. Compare this with the idea of a sink presented above. Clearly, asymptotic stability implies stability.

Clearly the idea of distance is important for the above ideas of stability. Sometimes, however, our normal Euclidean distance is not the best for analysing the stability of a critical point. Suppose a singularity is stable and orbits circle it in elliptical fashion, for instance. The distance to the critical point is changing, but remains bounded. It could be more convenient to find a notion of 'distance' in which the orbit is a fixed distance from the singularity. Also, with asymptotically stable points, if an orbit converges along some sort of 'elliptical spiral', its distance could, at some times, be moving away from the singularity, while overall converging to it. Again, a different distance would help.

Normally, distances are generalised with the study of metrics. In fact, we do not require a full metric here. The 'distance' functions we use are called Lyapounov functions.

DEFINITION 7.3 (Lyapounov Functions)

Let x_0 be a fixed point of the ODE $x' = f(x)$. Let W be some open set containing x_0. Then a differentiable function $V : W \to \mathbb{R}$ is a Lyapounov function if:

$$V(x_0) = 0 \ \ \& \ \ V(x) > 0 \ \ \forall x \in W \backslash \{x_0\}$$

$$\frac{\mathrm{d}V[\phi_t(x)]}{\mathrm{d}t} \leq 0 \ \ \forall x \in W \backslash x_0$$

Note that

$$\frac{\mathrm{d}V[\phi_t(x)]}{\mathrm{d}t} = DV_{\phi_t(x)} \frac{\mathrm{d}\phi_t(x)}{\mathrm{d}t} = DV_{\phi_t(x)} f[\phi_t(x)]$$

$V(t)$ is regarded as measuring the distance of the orbit from x_0 at time t; this is the reason for the requirement that $V(x) > 0$, as distances must be positive. The requirement $\frac{dV}{dt} \leq 0$ means that the distance is not increasing, which is what we desire for stability. Our normal Euclidean metric function, $V(x,y) = \sqrt{x^2+y^2}$, in 2D, is a valid Lyapounov function for some systems, if the orbits are circular, in the case of stability, or circular spirals, in the case of asymptotic stability. Other Lyapounov functions allow us to study systems which behave in less regular ways.

THEOREM 7.3 (Lyapounov Functions and Stability)

Let x_0 be a singularity of $x' = f(x)$. If there exists a Lyapounov function, then x_0 is stable and, for all x near x_0, $\phi_t(x)$ exists for all time. If the Lyapounov function has the additional property that $\frac{dV}{dt} < 0$, rather than simply ≤ 0, then x_0 is asymptotically stable.

PROOF:

Let W be the domain of the Lyapounov function. Pick $\varepsilon > 0$ such that $B_{2\varepsilon}(x_0) \subseteq W$. Let $m = \min_{x \in \partial B_\varepsilon(x_0)}\{V(x)\}$ (this exists by the extreme value theorem, as V is continuous and $\partial B_\varepsilon(x_0)$ is a compact set).

Now, $V(x_0) = 0$ and V is continuous, so $\exists \delta > 0 \ \ \forall x \in B_\delta(x_0) \ \ V(x) < m$. Then $B_\delta(x_0) \subseteq B_\varepsilon(x_0)$ and the supremum of V on $B_\delta(x_0)$ is less than the minimum value attained on $\partial B_\varepsilon(x_0)$.

Then pick any $x \in B_\delta(x_0)$. By the definition of a Lyapounov function, $V[\phi_t(x)] \leq V(x) \ \ \forall t > 0$ (since V is non-increasing). But $V(x) < m$, so we have $V[\phi_t(x)] < m \ \ \forall t > 0$. But m is the minimum value of V on the circle of radius ε, so $||\phi_t(x)|| < \varepsilon \ \ \forall t > 0$. So x_0 is stable.

We now prove the second part of the theorem; namely, that a strictly decreasing Lyapounov function implies asymptotic stability. We proceed by contradiction. Suppose there is no $\varepsilon > 0$ such that $x \in B_\varepsilon(x_0) \quad \Rightarrow \quad \phi_t(x) \to x_0$. Clearly, $\phi_t(x) \to x_0$ is equivalent to $V[\phi_t(x)] \to 0$; so our assumption is that this does not occur for any $x \neq x_0$.

Pick some x near x_0. Then $V(x)$ is decreasing and bounded below, so it converges to some ε, with $\varepsilon > 0$ by assumption. So, for all t, we have $\varepsilon \leq V[\phi_t(x)] \leq V(x)$. Let $\Omega = \{z : \varepsilon \leq V(z) \leq V(x)\}$. Orbits which begin in Ω stay there for all t. Ω is compact, so $\frac{dV}{dt}$ attains its maximum on Ω: $\max_{z\in\Omega}\{\dot{V}\} = -m < 0$. Since $\dot{V} \leq -m$ for all t, we have:

$$\begin{aligned} V[\phi_t(x)] &= V(x) + \int_0^t \frac{dV[\phi_s(x)]}{ds}\, ds \\ &\leq V(x) - mt \\ &\to -\infty \text{ as } t \to \infty \end{aligned}$$

But $V \geq 0$. Contradiction. ■

8 Dynamical Systems

In this section, we study the asymptotics of 2D (planar) systems of differential equations. We will tacitly assume that unique solutions exist for all time.

DEFINITION 8.1 (Limit Sets)

Let ϕ_t be the flow of a dynamical system. Take a point x. The <u>ω-limit set</u> of x, denoted $\omega(x)$, is the set of limit points of the curve $\phi_t(x)$ for $t \geq 0$. That is, $y \in \omega(x)$ iff there exists an increasing, unbounded sequence $\{t_n\}_{n=1}^{\infty}$ (tending to infinity) such that: $\phi_{t_n}(x) \to y$ as $n \to \infty$.

The simplest case is when an orbit converges to an asymptotically stable fixed point, x_0. In this case, $\omega(x) = \{x_0\}$. If the orbit is periodic, then for any point on the periodic solution, there is a sequence of times t_n at which the orbit returns to the point, so the periodic solution is the ω-limit set. If the orbit tends towards a periodic solution (as in the van der Pol equation, for instance), then for any point on the periodic solution, there is a sequence of times at which the orbit is near the given point, and tends to it as we progress through the sequence.

So, in summary, the ω-limit set is the set describes the eventual behaviour of the orbit; it is the set of points the solution tends towards.

DEFINITION 8.2 (Local Sections)

Pick a point $p \in \mathbb{R}^n$ and a smooth function $g : \mathbb{R}^n \to \mathbb{R}$, such that $g(p) = 0$ and $Dg(p) \neq 0$ (so, by the inverse function theorem, g is locally invertible near p). Let $S = \{x \in \mathbb{R}^n : g(x) = 0\}$ be the zero-level set of g. Now consider the system $x' = f(x)$ for some function $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $f(p) \neq 0$ and $Dg(p) \cdot f(p) \neq 0$, so that f is not orthogonal to the gradient of g; as the gradient of a function is always orthogonal to its level set, this means that f is not tangent to S at p.

Then we say that S is a local section at p. If S is a local section at x for all $x \in S$, we simlpy say S is a local section.

Roughly speaking, then, a local section at p is a curve that passes through p and is not tangent to $f(p)$ (and is locally invertible). This is clear enough from the name; a section is something which cuts (like a conic section is a shape made by cutting through a cone), so a local section cuts the orbit, meaning it must intersect it and not be tangent.

THEOREM 8.1 (Flow Box Theorem)

Let S be a local section at p and q some point such that $\phi_{t_0}(q) = p$ for some $t_0 > 0$. Then in some neighbourhood U of q, there is a function $\tau : U \to \mathbb{R} : q \mapsto t_0$ such that $\forall x \in U \;\; \phi_{\tau(x)}(x) \in S$. Roughly speaking, if we have a curve S passing through p, not tangent to the orbit through p, and a point q whose orbit passes through p, we can find a region around q and a function which finds the time points near q take to reach the curve S.

PROOF:

S is a local section, so it is a zero-level set of some function $g(x)$.

Let $G(x,t) = g[\phi_t(x)]$. Then $G(q,t_0) = g[\phi_{t_0}(q)] = g(p) = 0$. We see that:

$$\begin{aligned}\frac{\partial G}{\partial t}(q,t_0) &= Dg[\phi_{t_0}(q)] \left.\frac{\mathrm{d}\phi_t(q)}{\mathrm{d}t}\right|_{t_0} \\ &= Dg[\phi_{t_0}(q)] f[\phi_{t_0}(q)] \\ &= Dg(p) f(p) \\ &\neq 0\end{aligned}$$

So, as $G(q,t_0) = 0$ and $G'_t(q,t_0) \neq 0$, we see by the implicit function theorem that in some neighbourhood of q there exists a function τ such that $G(x,t) = 0$ iff $\tau(x) = t$. But $G(x,t) = 0$ iff $g[\phi_t(x)] = 0$, i.e., iff $\phi_t(x) \in S$. So τ is such that $\phi_{\tau(x)}(x) \in S$ and, in particular, $\tau(q) = t_0$.

■

THEOREM 8.2 (The Poincaré-Bendixson Theorem)

Consider an ODE in the plane. Take $x \in \mathbb{R}^2$ and let $\omega = \omega(x)$, the limit set of x, be non-empty, closed and bounded and not contain a singular point of the ODE. Then ω is a periodic orbit.

This means that, for a dynamical system in the plane, orbits must converge either to fixed points or to periodic solutions, or else they diverge to infinity. Effectively, this rules out chaotic behaviour for planar systems.